BEFORE THE
FEDERAL COMMUNICATIONS COMMISSION
WASHINGTON, D.C. 20554

In the Matter of Performance Measurements and Standards for Unbundled Network Elements and Interconnection	CC Docket No. 01-318
Performance Measurements and Reporting Requirements for Operations Support Systems, Interconnection, and Operator Services and Directory Assistance	CC Docket No. 98-56
Deployment of Wireline Services Offering Advanced Telecommunications Capability	CC Docket No. 98-147
Petition of Association for Local Telecommunications Services for Declaratory Ruling	CC Docket Nos. 98-147, 96-98, 98-141

COMMENTS OF CALIFORNIA

The People of the State of California and the California Public Utilities Commission ("California" or "CPUC") hereby file these initial comments in response to the Notice of Proposed Rulemaking ("NPRM") issued by the Federal Communications Commission ("FCC") on November 19, 2001 in the above-referenced proceedings. In this NPRM, the FCC seeks comment on whether it should adopt measurements and standards for evaluating the performance of incumbent local exchange carriers ("ILECs") in provisioning facilities used by competitive carriers in serving end use customers. The FCC specifically seeks comment in three major areas: (1) coordination between the FCC and the states in developing, applying and enforcing national performance measurements; (2) specific performance measurements and standards for pre-ordering, ordering, provisioning, and maintenance and repair; and (3) implementation and reporting requirements, and evaluation and statistical issues.

In these comments, California will comment on how state and federal regulation may be harmonized, the appropriate scope of performance standards, and the enforcement policies to effectuate these standards. California will also provide comment on the specific standards and measures proposed in the NPRM.

I. BACKGROUND

In decisions issued in 1999 and 2001, the CPUC adopted several provisions for a performance incentive plan in California. The final plan will consist of three fundamental components: performance measurement, performance assessment, and performance consequences. The first two components have been completed. The CPUC adopted OSS performance measurements in August 1999,[1] and revised them in May 2001.[2] In January, 2001, the CPUC adopted performance assessment

[1] Decision 99-08-020, August 5, 1999.

[2] Decision 01-05-087, May 24, 2001.

criteria (e.g., statistical tests and benchmarks).[3] A copy of each of these decisions is appended to these comments.

California is currently considering final provisions for a performance incentives plan that will monitor incumbent local exchange carriers' OSS performance to competitive local exchange carriers ("CLECs"). California anticipates adopting a final plan in the next month.. In addition to establishing the monetary amounts that will be paid for poor performance, the upcoming decision will make some final adjustments to the statistical model and benchmark criteria.

In light of the pendency of this matter before the CPUC, California is not in a position to provide comprehensive comments at this time. California will provide the FCC with its final decision on performance standards as soon as it becomes final.

II. COORDINATION OF FEDERAL AND STATE STANDARDS

In its NPRM, the FCC recognizes that a number of states have already adopted an extensive set of performance measurements, standards and penalty plans for the provisioning of UNEs, interconnection trunks, and collocation by ILECs. Most of the state effort has been in the context of section 271 proceedings. The FCC further recognizes that "[t]hese state efforts have been instrumental in evaluating and documenting incumbent LEC performance for purposes of

[3] Decision 01-01-037, January 18, 2001.

compliance with section 271's competitive checklist." NPRM, ¶ 15.

As discussed above, California has adopted a set of performance measurements and assessment methods, and on completion of the final phase, will implement a performance incentive plan to ensure that ILECs maintain markets that are open and that afford nondiscriminatory access to competing LECs.

California appreciates that the FCC seeks to "build on the states' pioneering efforts in developing performance measures and standards." NRPM, ¶ 15. At the same time, California understands the FCC's desire to adopt a set of national performance measures and standards. California believes that the best way to harmonize state and federal regulation is for the FCC to develop a minimum set of national performance measures and standards. These minimum measures and standards would serve three purposes: first, they would give states that have not yet adopted their own performance incentive plan a "headstart" in finalizing a state plan; second, for those states with limited resources that have not yet adopted their own measures and standards, minimum federal standards could be adopted in lieu of state standards; and third, minimum federal standards could define a threshold level of performance that, if met, will help foster competitive markets. At the same time, minimum national standards would allow states to continue implementing their own standards, to the extent that they exceeded the threshold set by national standards, tailored to the specific geographic markets in which they were developed.

In its NPRM, the FCC expresses concern that ILECs may be subject to a host of differing or conflicting state and federal standards, and that this could create an undue regulatory burden on the ILECs. California believes that such concern is overstated at this time. To date, only a handful of states have adopted a set of performance measures and standards for ILECs. Among these states, several share the same set of performance measures and standards for the ILEC. Thus, the number of differing or conflicting standards to which ILECs are subject currently is not large, and not unduly burdensome. The adoption of a minimum set of national standards will not increase the burden on ILECs, either because the ILECs are already subject to more stringent state standards, or because the federal standards will supplant less stringent state standards.[4]

California, however, acknowledges that over time, depending on how well state standards and minimum federal standards work together, it may make sense to adopt a single set of standards applicable to ILECs. Towards this end, California supports the idea of a state-federal joint board, created pursuant to section 410(b) of the Communications Act, that could monitor federal and state regulatory efforts, evaluate their efficacy, and, if appropriate, propose alternative frameworks. California, however believes that supplanting state measures and standards immediately, or even gradually, is premature at this time. Such state

[4] If the FCC adopts a set of standards like the ones suggested in the NPRM, that are substantially similar to standards already existing at the state level, the federal standards should pose no significant burden on the ILECs. Comparing the suggested federal measures with the California measures reveals much similarity.

measures and standards have only recently been developed, and should remain in place to determine their efficacy before the FCC, preferably through a federal-state joint board, considers a new set of measures and standards that are other than minimum national standards.

III. SCOPE

In its NPRM, the FCC seeks comment on whether national standards and measures should apply to all incumbent LECs, and to competitive LECs. California believes that performance standards and measures must apply to all incumbent LECs, since they are the carriers with bottleneck control of essential facilities and services necessary for competitors to access on a nondiscriminatory basis if truly competitive markets are to develop. Given the purpose of performance standards, it does not make sense to apply them to the competitive LECs.

California agrees that it may be appropriate to tailor the requirements for ILECs serving rural areas to reduce any undue burden on these carriers.

IV. ENFORCEMENT OF WHOLESALE OBLIGATIONS OF SECTION 271

In its NPRM, the FCC seeks comment on enforcement guidelines and policies when national performance standards and measures are not met. California supports a self-executing mechanism to ensure compliance with adopted standards and measures. California also believes that it is appropriate to presume competitive harm when violations occur. A case-by-case determination of competitive harm is not necessary, and will only mire the FCC and carriers in endless proceedings, at great time and expense.

In addition, California supports the imposition of the maximum monetary forfeitures available under federal law for violations. The FCC is correct in observing that penalties should not be so low such that ILECs are willing to violate the standards as a cost of doing business. Penalties set at the statutory maximum will best ensure that the measurements are seriously and conscientiously adhered to by the ILECs.

V. SPECIFIC MEASUREMENTS AND STANDARDS

California has adopted performance measures in the four performance areas listed in the NPRM: pre-ordering, ordering, provisioning, and maintenance/repair. Additionally, California has adopted performance measures in network performance, billing, database updates, collocation, and interfaces. In total, there are 44 performance measures. Table 1, set forth herein, lists the different measures in these areas.[5] While time does not allow us to enter each of these measures using the template presented in the NPRM, the CPUC's May 2001 Decision (D. 01-05-087) adopting these measures uses a very similar template. We are forwarding D. 01-05-087 with these comments to provide measurement summaries as well as more detailed information.

[5] Parties in the CPUC's proceeding have agreed that performance measures 8, 12, 13 and 22 sufficiently overlap with other measures to be excluded from incentive payments. The CPUC may adopt this exclusion in its final plan.

Table 1

California Performance Measures	
Pre-Ordering	
1	Average Response Time (to Pre-Order Queries)
Ordering	
2	Average FOC/LSC Notice Interval
3	Average Reject Notice Interval
4	Percentage of Flowthrough Orders
Provisioning	
5	Percentage of Orders Jeopardized
6	Average Jeopardy Notice Interval
7	Average Completed Interval
8	Percent Completed Within Standard Interval
9	Coordinated Customer Conversion as a Percentage On-Time
9a	Coordinated Frame Due Time Cutovers as a Percentage On-Time
10	PNP Network Provisioning
11	Percent of Due Dates Missed
12	Percent of Due Dates Missed Due to Lack of Facilities
13	Delay Order Interval to Completion Date (For Lack of Facilities)
14	Held Order Interval
15	Provisioning Trouble Reports (Prior to Service Order Completion)
15a	Average Time to Clear
16	Percentage Troubles in 30 Days for New Orders
17	Percentage Troubles in 10 Days for Non-Special Orders
18	Average Completion Notice Interval
Maintenance	
19	Customer Trouble Report Rate
20	Percentage of Customer Trouble Not Resolved Within Estimated Time
21	Average Time to Restore
22	POTS Out of Service Less Than 24 Hours
23	Frequency of Repeat Troubles in 30 Day Period
Network Performance	
24	Percent Blocking on Common Trunks
25	Percent Blocking on Interconnection Trunks
26	NXXs Loaded by LERG Effective Date
27	Network Outage Notification

Continued on next page.

Table 1 (continued)

Billing	
28	Usage Timeliness
29	Accuracy of Usage Feed
30	Wholesale Bill Timeliness
31	Usage Completeness
32	Recurring Charge Completeness
33	Non-Recurring Charge Completeness
34	Bill Accuracy
35	Timeliness of Billing Completion Notices
36	Accuracy of Mechanized Bill Feed
Database Updates	
37	Average Database Update Interval
38	Percent Database Accuracy
39	E911/911 MS Database Update Average
Collocation	
40	Average Time to Respond to a Collocation Request
41	Average Time to Provide a Collocation Arrangement
Interfaces	
42	Percentage of Time Interface is Available
43	Average Notification of Outages
44	Center Responsiveness

VI. SUNSET PROVISIONS

Since local markets, left alone, are natural monopolies, California believes that it does not make sense to "sunset" the provisions that monitor the health of competition and keep the ILEC from "backsliding" to its natural equilibrium, a monopoly. A true competitive local market will emerge only when competitors are able to employ new technologies that are not dependent on ILEC facilities and services. If and when that occurs, and competition is robust, will the need for performance incentives applicable to ILECs cease. In the meantime, enforcement of performance standards and measures must remain in effect in order to prevent the ILECs from returning to a natural monopoly. A worst case scenario would be where federal standards and enforcement would supplant state standards and enforcement, and then the federal standards would "sunset," unless states remain charged with enforcing federal standards. California is not aware of any time interval or "trigger" that could be used to signal that the usefulness of performance incentives is at an end.

Regarding federal standards that would not supplant state standards, we understand that a core set of minimum uniform national standards may be useful in making performance comparisons between states and companies. Currently, such standards could assist states as they calibrate and implement their new performance enforcement plans. As time passes, these measures and standards may help guide the maturation of quasi-competitive markets by allowing comparative evaluation of different performance enforcement schemes.

Respectfully submitted,

GARY M. COHEN
LIONEL B. WILSON
ELLEN S. LEVINE

By: /s/ Ellen S. LeVine

ELLEN LEVINE

505 Van Ness Ave.
San Francisco, CA 94102
Telephone: (415) 703-2047
Fax: (415) 703-2262

Attorneys for the People of the State of California and the California Public Utilities Commission

January 22, 2002

 Mailed 5/29/2001

Decision 01-05-087 May 24, 2001

BEFORE THE PUBLIC UTILITIES COMMISSION OF THE STATE OF CALIFORNIA

Order Instituting Rulemaking on the Commission's Own Motion into Monitoring Performance of Operations Support Systems.	Rulemaking 97-10-016 (Filed October 9, 1997)
Order Instituting Investigation on the Commission's Own Motion into Monitoring Performance of Operations Support Systems.	Investigation 97-10-017 (Filed October 9, 1997)

TABLE OF CONTENTS

O P I N I O N

Summary

Today we adopt revisions to the comprehensive framework for Operations Support Systems (OSS) performance measurements and standards that we adopted over a year ago in Decision (D.) 99-08-020.[1] These OSS measurements and standards are critical to ensuring that California's consumers have choices in local exchange telephone companies. OSS performance measurements and standards allow the Commission, the industry, and consumer advocates to measure and analyze the performance of Pacific and Verizon in providing their competitors nondiscriminatory access to their mechanized operating systems which store customer records and dispatch and monitor all network operations.

The revisions that we adopt today were proposed by Pacific, Verizon, and several of their major competitors (known as competitive local exchange carriers (CLECs)) after a comprehensive review of the OSS measurements, submeasurements, standards, and rules that we adopted last year in D.99-08-020. This group, collectively the Settling Parties, undertook the initial review of which OSS performance measurements and standards should be modified.[2] These are the companies providing or using OSS on a daily basis and therefore they have

---

[1] OSS are the manual and electronic systems by which competitive exchange carriers and the incumbent carriers, like Pacific Bell Telephone Company (Pacific) and Verizon California Inc. (Verizon, f/k/a GTE California, Inc.), exchange information regarding a number of logistical, technical, and administrative matters, including, but not limited to, billing, ordering, transfer of service, and new accounts.

[2] The Settling Parties are AT&T Communications of California, Inc. (AT&T), WorldCom, Inc. (WorldCom), Electric Lightwave, Inc. (ELI), ICG Access Services, Inc., Sprint Communications Company, L.P. (Sprint), Covad Communications Co. (Covad), Nextlink, Time Warner Telecom of California (TWTC), Pacific and Verizon.

the greatest knowledge and experience with Pacific's and Verizon's operating problems and capabilities. In addition to adopting major revisions to our OSS performance measurements and standards, we also adopt timetables for implementing the modifications and set a firm date to begin our 2001 review.

This decision does not address performance incentives for access to OSS subfunctions. On January 18, 2001, the Commission issued interim opinion D.01-01-037 in the incentive phase of this proceeding, which will establish remedies to ensure our OSS performance standards are met.

Although the parties agreed to significant modifications in the Joint Partial Settlement Agreement (JPSA) we adopt today, several issues regarding OSS performance measurements and standards remain in dispute. The Commission will address these issues in a later decision.

I. Procedural Background

On October 9, 1997, the Commission initiated this rulemaking proceeding as a procedural vehicle to accomplish the following three goals:

a. to determine reasonable standards of performance for Pacific and Verizon in their OSS;

b. to develop a mechanism that will allow the Commission to monitor improvements in the performance of OSS; and

c. to assess the best and fastest method of ensuring compliance if standards are not met or improvement is not shown.

In 1997, when the Commission initiated this proceeding, it recognized that it lacked the standards that it would need to evaluate Pacific's and Verizon's compliance with the requirements of the Telecommunications Act of 1996 (TA 96) and the Federal Communications Commission's (FCC) rules implementing TA 96. TA 96 requires incumbent local exchange carriers (ILECs) to provide competitors nondiscriminatory access to their operations support systems (OSS).[3]

The Commission also noted that this proceeding will prove critical to the Commission's ability to make an informed review of Pacific's OSS system under the § 271 application process of TA 96.[4] In August 1997, the FCC ruled that, with regard to those OSS subfunctions with retail analogs, a BOC must offer OSS subfunctions to CLECs that are on par with their own; they "must provide access to competing carriers that is equal to the level of access that the BOC provides to itself, its customers, or its affiliates, in terms of quality, accuracy, and timeliness."[5]

A "retail analog" exists when a BOC offers a retail service comparable to the one offered by a CLEC. When the BOC offers no comparable retail service, no retail analog exists. For those OSS sub-functions without retail analogs, a BOC must offer access sufficient to allow an efficient competitor "a meaningful

---

[3] See *In the Matter of Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, First Report and Order* (LCO), 12 FCC Rcd 15766, Paragraphs 516, 523.

[4] Regulators at the federal and state levels often allude to the "§ 271 process" and "§ 271 applications." They are referring to the statutory requirements under § 271 of the 1996 Telecommunications Act, which require Bell Operating Companies (BOCs) to open their local service markets to competition before being allowed to provide long distance services to their customers.

[5] See *In the Matter of Application of Ameritech Michigan Pursuant to Section 271 of the Communications Act of 1934, as amended, to Provide In-Region, InterLATA service in Michigan, Memorandum Opinion and Order*, 12 FCC Rcd 20543, 20618-19 [¶139] (1997) (Ameritech Opinion).

opportunity to compete."[6] The task of measuring progress towards these goals falls largely on state commissions.

On August 5, 1999 in D.99-08-020, the Commission adopted a comprehensive framework for OSS performance measurements and standards. In large part, the framework was the result of collaborative work among Pacific, Verizon, CLECs, and our Telecommunications Division staff. The Commission also adopted the parties' recommendation that the measurements and standards be reviewed and refined after six months. The "Joint Partial Settlement Agreement" (JPSA), the terms of which the Commission adopts today, grew out of this review process.

On March 24, 2000, pursuant to Rule 51.1(b) of the Commission's "Rules of Practice and Procedure," Pacific gave written notice to all parties of this proceeding that it would convene a settlement conference regarding the review of OSS performance measurements and standards. Following the initial settlement conference, interested parties met frequently over a six-month period to discuss revisions to the forty-four OSS measurements, and the many submeasures, standards, and business rules contained in the existing JPSA.

On July 18, 2000, the Settling Parties filed a "Joint Motion for Adoption of Partial Settlement Agreement Pursuant to Article 13.5 of the Commission's Rules of Practice and Procedure. On July 31, 2000, Verizon and Pacific filed separate motions in which they argued the merits of their positions on the "open" issues that remained among the Settling Parties. The CLEC members of the Settling

[6] See Ameritech Opinion, 12 FCC Rcd 20619 [¶ 141]. See also, BellSouth (Louisiana II) Opinion at ¶87 (citing Ameritech Opinion at 12 FCC Rcd at 20619).

Parties also filed a joint motion arguing that the Commission should adopt their collective positions regarding the open issues.

On July 31, 2000, NorthPoint Communications, Inc. (NorthPoint) and Rhythm Links, Inc. (Rhythms), neither of which joined the Settling Parties in the JPSA, filed comments on the settlement, the review process, and their position on open issues. In addition to presenting their position on open issues in these comments, NorthPoint and Rhythms argue that the review process is too long and burdensome for smaller competitors, particularly the data CLECs (DLECs); they recommend the Commission limit future reviews to one month.

On August 8, 2000, parties filed replies to the motions and comments. NorthPoint and Rhythms elected to forgo a reply brief and, instead, joined the CLECs in their reply brief. However, NorthPoint and Rhythms did not withdraw their proposal that the Commission limit the review process to a one month period and, therefore, did not join the CLECs' reply on that issue.

On August 17, 2000, the Office of Ratepayer Advocates (ORA) filed, pursuant to Rule 51.4 of the Commission's Rules of Practice and Procedure, comments in opposition to portions of the JPSA, recommending that proposed benchmarks for 16 measurements be established as parity measures before the Commission adopts the proposed settlement. In addition, ORA raised its concerns regarding the timeliness of its receipt of data.

On September 15, 2000, ORA filed a motion to withdraw its August 17th comments in exchange for the Settling Parties agreeing to give consideration to its concerns in the review. The Settling Parties filed a copy of the Memorandum of Understanding (MOU) that memorializes their agreement with ORA on September 20, 2000.

In addition, on November 6, 2000, the Settling Parties filed by motion a revised JPSA that expanded their July JPSA by adding aproximately 60 additional agreements. Finally, on February 13, 2001, Verizon, and three participating CLECs[7] filed a joint motion for approval of changes to Measurement 9. Verizon and the CLECs assert that their agreement resolves the disputed issue concerning Measurement 9.

II. The Revised Joint Partial Settlement Agreement

In their motion, the Settling Parties state that the JPSA represents their best efforts to ensure that OSS performance measurements and standards reflect the requirements of the real world. Towards this end, the Settling Parties have amended language, added two new measurements, deleted two measurements, included additional services and service levels, modified standards, clarified language, and agreed to meet and review OSS performance measurements again in March 2001. The Settling Parties have also proposed a timetable for implementing the changes entailed by adopting the JPSA.

In the JPSA, where the Settling Parties agreed about a proposed modification, the parties changed or added language to the standards we adopted in D.99-08-020. Where the parties disagreed about a proposed modification, they left the original language intact and recorded the proposed modification in an "open issues" document. The Settling Parties have also agreed to an implementation schedule for the JPSA, which they included under Section VIII of the JPSA. The November 6th proposed JPSA is attached at Appendix C.

[7] AT&T, WorldCom, and TWTC.

To facilitate our review of the JPSA, we summarize the purpose of each OSS performance measurement, identify the proposed modifications contained in the JPSA, and specify the disputed issues, referred to by the Settling Parties as "open issues." We provide this discussion in a separate appendix, Appendix B. We do this due to the length and technical nature of the summary.

III. Comments on the JPSA

The Settling Parties submit that the JPSA is reasonable in light of the whole record of competition in the California local exchange market, is consistent with the stated objectives of the Commission in this proceeding, and meets the Commission's public interest test for the approval of settlements. They assert that the measurements and standards of the JPSA are consistent with applicable law because they provide regulators with objective terms with which to measure the compliance of ILECs with TA 96. Furthermore, the JPSA, the Settling Parties observe, strikes a "reasonable compromise" between evaluating the ILECs' delivery of OSS and the administrative burdens of monitoring the ILECs' performance.

The Settling Parties also assert that the JPSA is in the public interest because many of the carriers that would be most directly affected by the standards by which Pacific and Verizon's OSS are provisioned have consented to its adoption. Because the CLECs who joined the Settling Parties will provide many local service options to California consumers, their concurrence in the JPSA, the Settling Parties collectively argue, makes the public's interest in the JPSA even greater.

NorthPoint and Rhythms participated in the February 2000 OSS performance measurement review but did not join the Settling Parties in signing the JPSA. On July 31, 2000, NorthPoint and Rhythms filed comments on the review process, on open issues, and on the proposed JPSA. On August 8, 2000, NorthPoint and Rhythms joined the CLEC members of the Settling Parties in filing a response to Pacific and Verizon on the open issues. Their positions on the open issues are reflected in Appendix B.[8] We discuss here their comments on the review process and adoption of the JPSA.

In their comments on the review process, NorthPoint and Rhythms state that only a very small group of CLECs were able to participate throughout the entire review process and, therefore, the proposed JPSA does not adequately represent the entire CLEC industry, especially the data CLECs' (DLECs)[9] interests. NorthPoint participated in the review process for approximately five weeks beginning in late May, and stated that during this period there were three day-long meetings at Pacific's offices in addition to three or more several-hour conference calls each week. During these meetings there were approximately 3-5 CLECs participating regularly and another 1 or 2 CLECs participating occasionally. NorthPoint decided not to sign the proposed JPSA because it was "unable to dedicate the resources needed to adequately address . . .[its]. . .concerns through this process without leaving an expansive list of open issues for the Commission to decide."

NorthPoint and Rhythms assert that most small and mid-sized CLECs do not possess the resources to effectively participate in an "almost 6 month

---

[8] Appendix B is a summary meant for informational purposes. The language of the revised JPSA, Appendix C, is controlling in the event that there are inconsistencies between the language of Appendices B and C.

[9] DLECs are those who only transport data traffic and do not transport voice communications.

non-stop process for reviewing these measures." They recommend that the Commission impose a review process that lasts no longer than one month in order to encourage broader CLEC participation.

While NorthPoint and Rhythms request the Commission change the review period proposed in the JPSA, they do not object to the Commission adopting all other portions of the JPSA. In their comments, they recognize the JPSA before us here is an improvement over the agreement we adopted in D.99-08-020, stating "the efforts of the CLECs that did participate throughout the entire process led to many improvements in the proposed JPSA."

On August 17, 2000, ORA filed comments pursuant to Rule 51.4 of the Commission's Rules of Practice and Procedure. In its comments, ORA objects to adoption of the JPSA because it relies on benchmarks rather than parity standards and because performance measurement data is not readily available to ORA. However, on September 15, 2000, after negotiating with the Settling Parties, ORA withdrew its Comments. In consideration for this, the Settling Parties agreed to undertake the following with respect to OSS performance measures:

- To include the Office of Ratepayer Advocates (ORA) staff in discussions about the functionality of the OSS performance measures website and the configuration of the performance data on the website, and
- In the context of the March 2001 annual review of OSS performance measures, to consider amending the standards of at least five performance measures, which are currently benchmark standards, to either a party standard or standard based upon historical data.

IV. The Revised JPSA is Reasonable, Consistent with the Law, and in the Public Interest

A. Summary

Rule 51.1 of the Commission's "Compiled Rules of Practice and Procedure" governs the proposal of settlements. Rule 51.1(e) requires that a settlement be "reasonable in light of the whole record, consistent with law, and in the public interest" before it is approved. Based on the discussion here, we find that the JPSA is reasonable in light of the whole record, consistent with law, and in the public interest. Therefore, we will adopt the agreement.

B. Discussion

The JPSA is the result of lengthy negotiations among Pacific, Verizon, and several CLECs. The Settling Parties reviewed all of the measurements and standards that were adopted by the Commission in D.99-08-020. They also reviewed those issues that the Commission specifically required parties to re-negotiate in the August 1999 decision.

The "open issues" on which the Settling Parties cannot agree have been discussed extensively in the motions and replies submitted by the parties. Because some of the open issues involve further modifications to the measurements and standards that we adopted in D.99-08-020, the JPSA should be received as a partial statement of OSS performance standards and measurements. We have indicated in Appendix B which elements are subject to revision, pending our resolution of the open issues.

As a threshold matter, the Settling Parties seek to limit the application of the JPSA. "By seeking approval of the JPSA, the Settling Parties make no representation that the JPSA constitutes a definitive or a conclusive standard for Pacific's or GTE's compliance with the Telecommunications Act of 1996."

Furthermore, AT&T reserves its rights to argue that "parity, not benchmarks, are the appropriate performance measures under applicable law." Still further, by agreeing to the terms of the JPSA, Pacific and Verizon make no commitments or admissions regarding the "propriety or reasonableness of establishing performance remedies."

The limitation the Settling Parties place on the JPSA are consistent with the evolving process the Commission is using to develop and implement OSS performance measurements. The JPSA before us today is more comprehensive than the JPSA we approved in D.99-08-020. As the Settling Parties observe, the JPSA "embodies the best efforts of the CLECs, Pacific, and GTE to modify, as necessary or appropriate, the performance measurements approved by the Commission in D.99-08-020." We will be refining the measurements when we decide the open issues and the Settling Parties themselves propose reviewing the measurements again in March 2001. The performance measurements are only one measure of compliance with TA 96, and therefore, by approving the JPSA we are not concluding that it represents a definitive or conclusive standard for Pacific's or Verizon's compliance with TA 96.

The Settling Parties have submitted a document clearly outlining the specific elements of their proposed changes along with the rationale for their modifications to the measurements, standards, and business rules we adopted in D.99-08-020. While we adopt the revised JPSA based on our own independent analysis, we note that the JPSA represents the consensus among fiercely competitive parties that normally agree on very little.

We find that the JPSA is a proposal that provides a comprehensive update to the OSS performance measurements and standards we adopted in D.99-08-020. The JPSA adds new services, service levels, and products, includes two new measurements, deletes one service measurement because a quicker alternative is available, and clarifies existing business rules. The proposal reflects the experience that industry participants have gained since our earlier proceeding and provides substantial progress toward fully achieving our goal to provide competitors nondiscriminatory access to Pacific's and Verizon's OSS. The JPSA articulates in a detailed manner the very categories by which the Commission, the industry, and consumer advocates can measure, analyze, and review the success of Pacific and Verizon in providing nondiscriminatory access to OSS.

Promoting competition in California's local exchange telephone market, as required by TA 96 and California Pub. Util. Code §§ 709.5 and 709.7 is a significant public policy goal of this Commission. To achieve our goal, competitors must have access to pre-ordering, ordering, provisioning, maintenance and network performance, database updates, collocation, and interface information (the OSS subfunctions) from Pacific and Verizon that is equal to the level of access in terms of quality, accuracy, and timeliness that Pacific and Verizon provide themselves, their customers, and their affiliates. Without this nondiscriminatory access, competitors that need to use Pacific and Verizon's network to provide local exchange service cannot provide their customers quality service. Therefore, the revised JPSA is reasonable and in the public interest.

The JPSA is consistent with applicable law because it offers a system of objective terms by which the Commission can measure, discuss, and analyze the success of Pacific and Verizon in meeting their legal duties under TA 96 and the FCC rules implementing the 1996 Act. The measurements and standards contained in the JPSA will greatly assist the Commission in making legal and factual judgments about OSS subfunctions both when we review any current or future Section 271 applications by Pacific and also when we review facts in connection with OSS performance incentives.

NorthPoint and Rhythms request the Commission change the review procedures contained in Section VI of the JPSA. In Section VI, the Settling Parties agree to reconvene on or around March 1, 2001 to review the effectiveness of and modifications to the performance measurements approved by the Commission in this proceeding. The parties agree to conclude this review within 90 days of its commencement and to submit their revisions to the Commission, together with any disputed issues, within the 90-day review period. NorthPoint and Rhythms request we shorten this review period to 30 days in order to ensure that smaller CLECs can fully participate in the process.

The Settling Parties spent six months in reviewing and negotiating the proposed JPSA. Their agreement to limit the review period in 2001 appears to be an accommodation to NorthPoint's and Rhythm's concern. We have found it very beneficial for the parties to spend considerable time and effort identifying and discussing the very detailed and technically complex OSS issues involved in setting OSS performance measurements and standards. Without the parties doing this work, the Commission would not have the comprehensive OSS measurements and standards it has today. Both NorthPoint and Rhythms were able to participate in portions of this review process and other DLECs can also identify specific areas of interest and participate in those areas of review. We find the JPSA's three-month review period to be reasonable and, therefore, adopt it.

A final issue that the Settling Parties bring before us in the JPSA is their objection to the inclusion of Commission ordered language in the actual settlement document. In D.99-08-020, the Commission decided the disputed issues before it and inserted our requirements directly into the proposed JPSA format, making Appendix B of the decision a complete list of all adopted OSS measurements, standards, auditing, reporting, implementation, and review procedures. In the proposed JPSA before us today, the Settling Parties have deleted the Commission-added language from the statement of OSS measurements and standards because they believe inclusion in the proposed JPSA of this language creates an invalid impression that the parties themselves have reached an agreement on these measurements.

The Settling Parties "expressly agree" that any language added by the Commission in its D.99-08-020 decision which obligates Pacific or Verizon "to provide certain types of OSS access or to perform certain auditing or reporting requirements remains enforceable as part of that decision and is not rendered unenforceable as a result of having been removed by the parties." Nevertheless, the Settling Parties request that, in the future, the Commission avoid adding such language to the JPSA. The Settling Parties propose that the Commission include such language with the ordering paragraphs of the decision by which the Commission adopts the JPSA.

We should accommodate the Settling Parties request to not include our modifications directly in their signed settlement document. However, we do not agree with the Settling Parties that the Commission's modifications should only be contained in the ordering paragraphs of its decisions. We find it beneficial to have all OSS performance measurements and standards available in one place for ease of reference and to ensure the public and all interested parties are fully informed.

Therefore, we should include at Appendix C a separate listing of the Commission modifications in D.99-08-020 together with the JPSA we adopt today. The Settling Parties have facilitated this process by placing the Commission's D.99-08-020 adopted language at the front of their revised JPSA. This addition is clearly identified as the work of the Commission. This supplement and the revised JPSA, together, will serve as a single statement of our adopted OSS performance measurements and standards.

C. Next Steps

The Commission will resolve the open issues and then schedule a prehearing conference to begin the 2001-review process.

V. Comments on Draft Decision

The draft decision of Administrative Law Judge Walwyn in this matter was mailed to the parties in accordance with Section 311(g)(1) of the Pub. Util. Code and Rule 77.7 of the Rules of Practice and Procedure. Comments were filed by AT&T, XO, WorldCom, and Pacific (Joint Commenters) and Verizon on May 14, 2001, and reply comments were filed on May 21, 2001.

We adopt the technical corrections recommended by the Joint Commenters and the recommendation of Joint Commenters and Verizon to resolve the open issues before beginning the 2001 review. While September 2001 is a reasonable timeframe for resolving the open issues, our resources and other priorities do not allow us to commit to the specific schedules requested by commenters.

Findings of Fact

1. On August 5, 1999, the Commission adopted a comprehensive framework for OSS performance measurements and standards, which was largely the result of collaborative work among Pacific, Verizon, CLECs, and our Telecommunications staff.

2. On July 18, 2000, several California CLECs and ILECs, the Settling Parties, filed a "Joint Motion for Adoption of Partial Settlement Agreement Pursuant to

Article 13.5 of the Commission's Rules of Practice and Procedure." The Settling Parties later added further agreements to the JPSA and submitted the revisions to the Commission by motions on November 6, 2000 and February 13, 2001.

3. Several proposals to make additional modifications to the JPSA remain in dispute among the Settling Parties, NorthPoint, and Rhythms.

4. On August 17, 2000, ORA filed, pursuant to Rule 51.4 of the Commission's Rules of Practice and Procedure, comments opposing portions of the revised JPSA.

5. On September 20, 2000, the Settling Parties filed "Response of Settling Parties to the Office of Ratepayer Advocates' Motion to Withdraw Comment: Confirmation of Resolution of Issues." ORA and the Settling Parties have entered into an MOU in which the Settling Parties agree to address some of ORA's comments in the 2001 review of OSS performance measurements and standards.

6. The revised JPSA articulates in a detailed manner the very categories by which the Commission, the industry, and consumer advocates can measure, analyze, and review the success of Pacific and Verizon in providing nondiscriminatory access to OSS.

7. The revised JPSA adds new services, service levels, and products, includes two new measurements, deletes two service measurements because a quicker alternative is available, and clarifies existing business rules.

8. The OSS performance measurements and standards set forth in the revised JPSA provide a critical framework within which the Commission can assess the ILECs' compliance with the Telecommunications Act of 1996, and their delivery of nondiscriminatory OSS services. The OSS performance and standards

outlined in the revised JPSA will also prove critical in the 271 application process for Pacific.

Conclusions of Law

1. The revised JPSA is a proposal that provides a comprehensive update to the OSS performance measurements and standards we adopted in Decision (D.) 99-08-020.

2. The revised JPSA reflects the experience that industry participants have gained since we issued D.99-08-020.

3. The revised JPSA's proposal of a three-month initial review process among interested parties is reasonable.

4. The revised JPSA submitted by the Settling Parties is reasonable in light of the whole record, consistent with law, and in the public interest.

5. The issues remaining in dispute, the open issues, are identified at Appendix B and should be addressed in a later Commission decision.

6. The Memorandum of Understanding between ORA and the Settling Parties should be addressed by the Settling Parties in the 2001 review of OSS performance measurements and standards.

7. The language which the Commission adopted as revisions to the JPSA in D.99-08-020, together with the November 6, 2000 revised JPSA and the February 13, 2001 Verizon and participating CLECs Measurement 9 agreement, constitute our adopted framework for OSS performance measurements and standards in California. The revised JPSA should be considered a partial statement of OSS performance measurements and standards since disputed issues remain such that the resolution of those issues, identified at Appendix B, place portions of the revised JPSA subject to amendment.

O R D E R

IT IS ORDERED that:

1. We adopt the revised JPSA at Appendix C.

2. The open issues identified by parties, and summarized in Appendix B, shall be addressed in a future decision.

3. The schedule for the 2001 Operations Support Systems performance measurements review shall be set by separate ruling.

This order is effective today.

Dated May 24, 2001, at San Francisco, California.

LORETTA M. LYNCH
President
HENRY M. DUQUE
RICHARD A. BILAS
CARL W. WOOD
GEOFFREY F. BROWN
Commissioners

APPENDIX A

************** SERVICE LIST *************
Last Update on 18-APR-2001 by: LIL
R9710016 LIST
I9710017

************** APPEARANCES **************

Randolph Deutsch
Attorney At Law
AT&T COMMUNICATIONS OF CALIFORNIA, INC.
795 FOLSOM STREET
SAN FRANCISCO CA 94107
(415) 442-5560
deutsch@att.com

Gregory H. Hoffman
AT&T CORP.
ROOM 2161
795 FOLSOM STREET
SAN FRANCISCO CA 94107
(415) 442-3776
ghhoffman@att.com
For: AT&T

Jeffrey F. Beck
JILLISA BRONFMAN
BECK & ACKERMAN
4 EMBARCADERO CENTER, SUITE 760
SAN FRANCISCO CA 94111
(415) 263-7300
smalllecs@aol.com

Anita C. Taff-Rice
Attorney At Law
BLUMENFELD & COHEN
4 EMBARCADERO CENTER, SUITE 1170
SAN FRANCISCO CA 94111
(415) 394-7500
anitataffrice@earthlink.net
For: Rhythms Links, Inc.

Stephen P. Bowen
CHRISTINE A. MAILLOUX
Attorney At Law
BLUMENFELD & COHEN
4 EMBARCADERO CENTER, SUITE 1170
SAN FRANCISCO CA 94111
(415) 394-7500
stevebowen@earthlink.net

Lesla Lehtonen
Staff Attorney
CALIFORNIA CABLE TELEVISION ASSN.
PO BOX 11080
4341 PIEDMONT AVENUE
OAKLAND CA 94611-0080
(510) 428-2225
darlene@calcable.org

Barbara L. Snider
Attorney At Law
CITIZENS COMMUNICATIONS
9324 W. STOCKTON BLVD., SUITE 100
ELK GROVE CA 95758
(916) 691-5550
bsnider@czn.com

Charles E. Born
Director, Regulatory Affairs
CITIZENS COMMUNICATIONS
PO BOX 340
ELK GROVE CA 95759
(916) 691-5550
cborn@czn.com

John R. Gutierrez
COLE, RAYWID & BRAVERMAN, LLP
895 B STREET, NO. 526
HAYWARD CA 94541
(510) 881-2281
jgutierrez@crblaw.com
For: PAC-WEST TELECOMM, INC.

Mark Schreiber
E.GARTH BLACK / MARCK P.SCHREIBER
Attorney At Law
COOPER, WHITE & COOPER, LLP
201 CALIFORNIA STREET, 17TH FLOOR
SAN FRANCISCO CA 94111
(415) 433-1900
mschreiber@cwclaw.com

Margaret Robbins
Asst. Gneral Counsel
COVAD COMMUNICATIONS CO.
4250 BURTON DRIVE
SANTA CLARA CA 95054
(408) 987-1105
mrobbins@covad.com
For: Covad Communications

Treg Tremont
Attorney At Law
DAVIS WRIGHT TREMAINE
ONE EMBARCADERO CENTER, SUITE 600
SAN FRANCISCO CA 94111-3834
(415) 276-6500
tregtremont@dwt.com
For: XO and New Edge Network, Inc.

Jane Whang
Attorney At Law
DAVIS WRIGHT TREMAINE LLP
ONE EMBARCADERO CENTER, SUITE 600
SAN FRANCISCO CA 94111-3834
(415) 276-6500
janewhang@dwt.com

Peder Gunderson
ELECTRIC LIGHTWAVE, INC.
7700 NE 77TH AVENUE
VANCOUVER WA 98662-0606
(360) 816-5001
pgunderson@eli.com

GENESIS COMMUNICATIONS
12750 HIGH BLUFF DRIVE, SUITE 200
SAN DIEGO CA 92130-2083
(858) 792-2400

John Clark
Attorney At Law
GOODIN MACBRIDE SQUERI RITCHIE & DAY LLP
505 SANSOME STREET, 9TH FLOOR
SAN FRANCISCO CA 94111
(415) 765-8443
jclark@gmssr.com

Cheryl Ann Klepper
GTE CALIFORNIA, INCORPORATED
ONE GTE PLACE, M.C. CA500LB
THOUSAND OAKS CA 91362-3811
(805) 372-8333
cheryl.klepper@telops.gte.com
For: GTE California, Incorporated

Janice L. Grau
Administrative Law Judge Division
RM. 5011
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1223
jlg@cpuc.ca.gov
For: Office of RatepayerAdvocates

Marlin Ard
Attorney At Law
HERSHNER HUNTER ANDREWS NEILL&SMITH, LLP
SODA CREEK GALLERY BLDG
178 S. ELM, ROOM 205
SISTERS OR 97759
(805) 372-6000
maratty@coinet.com

Cheryl Hills
ICG TELECOM GROUP, INC.
180 GRAND AVENUE, SUITE 800
OAKLAND CA 94612
(510) 251-7033

Ira Kalinsky
Legal Division
RM. 5027
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2130
kal@cpuc.ca.gov

Earl Nicholas Selby
Attorney At Law
LAW OFFICES OF EARL NICHOLAS SELBY
418 FLORENCE STREET
PALO ALTO CA 94301-1705
(650) 323-0990
ens@loens.com

Glenn Harris
Attorney At Law
LAW OFFICES OF EARL NICHOLAS SELBY
418 FLORENCE STREET
PALO ALTO CA 94301
(650) 323-0990
gaharris@wenet.net
For: ICG TELECOM GROUP, INC.

Andrew Ulmer
Attorney At Law
MBV LAW, LLP
101 VALLEJO ST.
SAN FRANCISCO CA 94111
(415) 781-4400
andrew@mbvlaw.com
For: California ISP Association ,Inc.

Gregory L. Bowling
MCCUTCHEN,DOYLE,BROWN & ENERSEN, LLP
THREE EMBARACDERO CENTER
SAN FRANCISCO CA 94111-4067
(415) 393-2000

Judith A. Holiber
MORGENSTEIN & JUBELIRER LLP
ONE MARKET PLAZA
SPEAR STREET TOWER, 32ND FLOOR
SAN FRANCISCO CA 94105

For: GTE

Mary E. Wand
Attorney At Law
MORRISON & FOERSTER LLP
425 MARKET STREET
SAN FRANCISCO CA 94105
(415) 268-7201
mwand@mofo.com

Theresa Cabral
MORRISON & FOERSTER, LLP
101 YGNACIO VALLEY ROAD, SUITE 450
WALNUT CREEK CA 94596
(925) 295-3370
tcabral@mofo.com
For: PacWest Telecomm, Inc.

Helen M. Mickiewicz
Legal Division
RM. 5123
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1319
hmm@cpuc.ca.gov

Penny Bewick
NEW EDGE NETWORKS, INC.
PO BOX 5159
3000 COLUMBIA HOUSE BLVD., 106
VANCOUVER WA 98668
(360) 906-9775
pbewick@newedgenetworks.com

Karen M. Potkul
Attorney At Law
NEXTLINK CALIFORNIA, INC.
1924 E. DEERE AVENUE, STE 110
SANTA ANA CA 92705
(949) 417-7766
kpotkul@nextlink.net

Steven Gorosh
Vice President & General Counsel
NORTHPOINT COMMUNICATIONS
303 SECOND STREET
SAN FRANCISCO CA 94107
(415) 403-4003

Martin A. Mattes
Attorney At Law
NOSSAMAN GUTHNER KNOX & ELLIOTT, LLP
50 CALIFORNIA STREET, 34TH FLOOR
SAN FRANCISCO CA 94111-4799
(415) 398-3600

Cindy Camillucci
Marketing Manager
OMNICOM "CLEAR AROUND THE WORLD"
SUITE 340
1919 WILLIAMS STREET
SIMI VALLEY CA 93065
(805) 520-4047
ccamillucci@its-omni.com

Robert J. Gloistein
Attorney At Law
ORRICK HERRINGTON & SUTCLIFFE, LLP
400 SANSOME STREET
SAN FRANCISCO CA 94111-3143
(415) 773-5900
rgloistein@orrick.com

Fredrick Gamble
Consultant
OUTER SPHERE, INC.
1445 CEDAR AVENUE
SAN LEANDRO CA 94579
(510) 351-0518
fredgamble@earthlink.net
For: COX TELCOM CALIFORNIA, LLC

Yvonne Gamble
Consultant
OUTER SPHERE, INC.
1445 CEDAR AVENUE
SAN LEANDRO CA 94579
(510) 351-0518
yvonnegamble@outersphere.com
For: COX TELCOM CALIFORNIA, LLC

Robert Townsend
PAC-WEST TELECOMM, INC.
1776 W. MARCH LANE, SUITE 250
STOCKTON CA 95207
(209) 926-3032
rtownsen@pacwest.com
For: PAC-WEST TELECOMM, INC.

Ed Kolto Wininger
Attorney At Law
PACIFIC BELL
140 NEW MONTGOMERY ST., ROOM 1619
SAN FRANCISCO CA 94105
(415) 545-9422
ed.kolto.wininger@pactel.com

mmattes@nossaman.com

Hugh Osborne
PACIFIC BELL
ROOM 1323
140 NEW MONTGOMERY STREET
SAN FRANCISCO CA 94105

Peter A. Casciato
Attorney At Law
PETER A. CASCIATO, PC
8 CALIFORNIA STREET, SUITE 701
SAN FRANCISCO CA 94111-4825
(415) 291-8661
pcasciato@telocity.com

Steve Kukta
SPRINT COMMUNICATIONS COMPANY LP
100 SPEAR STREET, SUITE 930
SAN FRANCISCO CA 94105
(415) 371-7178
stephen.h.kukta@mail.sprint.com

Andrew O. Isar
Director, Industry Relations
TELECOMMUNICATIONS RESELLERS ASSN.
4312 92ND AVENUE, N.W.
GIG HARBOR WA 98335
(253) 265-3910
aisar@harbor-group.com

Regina Costa
Telecommunications Research Director
THE UTILITY REFORM NETWORK
711 VAN NESS AVENUE, SUITE 350
SAN FRANCISCO CA 94102
(415) 929-8879 X312
rcosta@turn.org

Elaine M. Duncan
SUSAN D. ROSSI, ESQ.
Attorney At Law
VERIZON CALIFORNIA INC.
711 VAN NESS AVENUE, SUITE 300
SAN FRANCISCO CA 94102
(415) 474-0468
elaine.duncan@verizon.com

Evelyn C. Lee
Attorney At Law
WORLDCOM, INC.
201 SPEAR STREET, 9TH FLOOR
SAN FRANCISCO CA 94105
(415) 228-1264
evelyn.lee@wcom.com

*********** STATE EMPLOYEE ***********

Michael C. Amato
Telecommunications Division
RM. 3203
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1863
mca@cpuc.ca.gov

Richard A. Bilas
Executive Division
RM. 5218
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-3703
rb1@cpuc.ca.gov

Natalie Billingsley
Office of Ratepayer Advocates
RM. 4101
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1368
nxb@cpuc.ca.gov

William M Braxton
Telecommunications Division
AREA 3-D
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-4153
bil@cpuc.ca.gov

Brian M. Chang
Office of Ratepayer Advocates
RM. 4205
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1333
bmc@cpuc.ca.gov

Cherrie Conner
Telecommunications Division
AREA 3-D
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2767
chr@cpuc.ca.gov

For: WorldCom, Inc.

Linda L. Gustafson
Office of Ratepayer Advocates
RM. 4102
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2385
llg@cpuc.ca.gov

Karen Jones
Administrative Law Judge Division
RM. 5041
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1083
kaj@cpuc.ca.gov

Ramesh Joshi
Office of Ratepayer Advocates
RM. 4205
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1842
ruj@cpuc.ca.gov

Paul King
Telecommunications Division
AREA 3-E
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2338
pwk@cpuc.ca.gov

Jonathan Lakritz
Executive Division
AREA 3-E
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2642
jol@cpuc.ca.gov

Yasugi Otsuka
NEVADA PUBLIC UTILITIES COMMISSION
1150 E. WILLIAM STREET
CARSON CITY NV 89701-3109

Nazmeen Rahman
Telecommunications Division
AREA 3-D
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1625
nar@cpuc.ca.gov
For: TELECOM DIVISION

Jacqueline A. Reed
Administrative Law Judge Division
RM. 5117
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2935
jar@cpuc.ca.gov

Jan Reid
Office of Ratepayer Advocates
RM. 4209
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-1546
ljr@cpuc.ca.gov

Randy Chinn
SENATE ENERGY UTILITIES & COMMUNICATIONS
STATE CAPITOL, ROOM 4040
SACRAMENTO CA 95814
(916) 445-9764
randy.chinn@sen.ca.gov

Maria E. Stevens
Executive Division
RM. 5109
320 WEST 4TH STREET SUITE 500
Los Angeles CA 90013
(213) 576-7012
mer@cpuc.ca.gov

Jensen Uchida
Energy Division
AREA 4-A
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-5484
jmu@cpuc.ca.gov
For: CPUC Telecom Division

Christine M. Walwyn
Administrative Law Judge Division
RM. 5101
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2301
cmw@cpuc.ca.gov

Lester Wong
Executive Division
RM. 5216
505 VAN NESS AVE
San Francisco CA 94102

Ting-Pong Yuen
Telecommunications Division
AREA 3-D
505 VAN NESS AVE
San Francisco CA 94102
(415) 703-2895
tpy@cpuc.ca.gov

*********** INFORMATION ONLY ***********

Bonnie K. Alexander
ALEXANDER CONSULTING
4944 CORTE PLAYA DE CASTILLA
SAN DIEGO CA 92124-1785
(858) 573-1785
bonnie@adnc.com

Candace Clark
CAP GEMINI TELECOMMUNICATIONS
801 E. CAMPBELL ROAD, SUITE 475
RICHARDSON TX 75081
(972) 761-9900
canclark@usa.capgemini.com

Jerry Stroud
CAPGEMINI COMMUNICATIONS
100 WALNUT AVENUE
CLARK NJ 07066
(732) 382-5400
jstroud@usa.capgemini.com

Heidi Sieck Williamson
Dept Of Telecommunications & Information
CITY & COUNTY OF SAN FRANCISCO
875 STEVENSON STREET, 5TH FLOOR
SAN FRANCISCO CA 94103
(415) 554-0811
heidi_sieck-williamson@ci.sf.ca.us

Thomas J. Burke
CONTEL SERVICE CORPORATION
#CA500 GCF
ONE GTE PLACE
THOUSAND OAKS CA 91362-3811

Bernard H. Chao
Esquire
COVAD COMMUNICATIONS COMPANY
2330 CENTRAL EXPRESSWAY
SANTA CLARA CA 95050
(408) 844-7500
bchao@covad.com

(415) 703-3702
llj@cpuc.ca.gov

Christopher Goodpastor
Asst. General Counsel
COVAD COMMUNICATIONS COMPANY
816 CONGRESS AVENUE, SUITE 1100
AUSTIN TX 78701
(512) 493-5777
cgoodpas@covad.com

Marilyn Mcalister
Ilec Relations Manager
COVAD COMMUNICATIONS COMPANY
2330 CENTRAL EXPRESSWAY
SANTA CLARA CA 95050
(408) 844-7500
mcalister@covad.com

Richard Smith , Jose Jimenez
COX CALIFORNIA TELCOM, LLC, DBA COX COMM
2200 POWELL STREET, SUITE 795
EMERYVILLE CA 94608-2618
(510) 923-6220
richard.smith@cox.com

Juanita Harris
DEPARTMENT OF JUSTICE/TELECOM TASK FORCE
ANTITRUST DIVISION NO.8104
1401 H STREET NW, STE 8000
WASHINGTON DC 20530
(202) 514-5642
juanita.harris@usdoj.gov

Frederick R. Duda
Commissioner Emeritus
4034 BRADBURY DRIVE
MARIETTA GA 30062-6135
(707) 321-4400
fbduda@juno.com

Jo Ann G. Hill
Sr. Manager, Public Policy
FIRST WORLD COMMUNICATIONS
222 SOUTH HARBOR BLVD., SUITE 400
ANAHEIM CA 92805
(858) 552-8010

Margo Friedrich
Regulatory Affairs
GTE CALIFORNIA INCORPORATED
711 VAN NESS AVE., STE 300
SAN FRANCISCO CA 94102
margo.friedrich@telops.gte.com

Cynthia Walker
Dir. Of Governmental & External Affairs
ICG TELECOM GROUP, INC.
180 GRAND AVENUE, SUITE 800
OAKLAND CA 94612
(510) 239-7089
cynthia_walker@icgcomm.com

Erin R. Swansiger
KELLEY DRYE & WARREN, LLP
1200 19TH STREET, N.W., SUITE 500
WASHINGTON DC 20036
(202) 887-1232
eswansiger@kelleydrye.com
For: KELLEY DRYE & WARREN LLP

Laura E. Arsaga
Associate
LAW OFFICES OF E. NICHOLAS SELBY
418 FLORENCE
PALO ALTO CA 94301
(650) 323-0990
learsaga@wenet.net

Kim Logue
Regulatory Analyst
LCI INTERNATIONAL TELECOM CORP.
4250 N. FAIRFAX DRIVE, 12W002
ARLINGTON VA 22203
(703) 363-4321
kim.logue@qwest.net

David Simpson
MANDEL, BUDER & VERGES
101 VALLEJO STREET
SAN FRANCISCO CA 94111
(415) 781-4400
das@bizlaw.net

Tracey Buck-Walsh
Regulatory Vp, Western Region
MPOWER COMMUNICATIONS CORP.
2535 CAPITOL OAKS DRIVE, SUITE 250
SACRAMENTO CA 95833

Esther Northrup
Attorney At Law
NEXTLINK CALIFORNIA, INC.
5771 COPELY DRIVE
SAN DIEGO CA 92111
(858) 244-4910

Doug Hsiao
Assistant General Counsel
RHYTHMS LINKS, INC.
9100 E MINERAL CIRCLE
ENGLEWOOD CO 80112

Renee Van Dieen
SPRINT COMMUNICATIONS COMPANY LP
7TH FLOOR
1850 GATEWAY DRIVE
SAN MATEO CA 94404

Morton J. Posner
MICHAEL ROMANO
SWIDLER BERLIN SHEREFF FRIEDMAN, LLP
3000 K STREET, N.W., SUITE 300
WASHINGTON DC 20007

TELE-MATIC CORPORATION
SUITE 100
67 INVERNESS DR. EAST
ENGELWOOD CO 80112-5117

Melissa Waksman
Attorney At Law
XO CALIFORNIA, INC.
ONE FRONT STREET, SUITE 1850
SAN FRANCISCO CA 94111
(415) 901-3816
melissa.waksman@xo.com

ehnorthrup@nextlink.com

(END OF APPENDIX A)

APPENDIX B

APPENDIX B
Page 1

APPENDIX B - SUMMARY OF CHANGES TO OSS PERFORMANCE MEASUREMENTS CONTAINED IN THE NOVEMBER 6, 2000 JOINT PARTIAL SETTLEMENT AGREEMENT (JPSA) AND DISPUTED ISSUES REMAINING FOR COMMISSION RESOLUTION

A. Pre-Ordering Measurements.

Measure 1: Average Response Time (to Pre-Order Queries).

This measurement calculates the average time that it takes Pacific/Verizon to respond to pre-order queries. CLECs submit pre-order queries to Pacific/Verizon to determine the availability of services requested by the customer, to verify customer information (including which services the customer is currently receiving) to request a due date for a service appointment, etc. The measurement requires separate reporting based on the type of information requested. The time it takes for the CLEC to obtain a response to these queries, often while the customer is on the line, has an important effect on how the customer perceives the CLEC's capabilities.

The Settling Parties propose modifying the description of this measurement to include language regarding the inclusion of loop qualifications. They offer a new formula for calculating this measurement which reflects their agreement on measurable standards. The Settling Parties propose amending the measurable standard regarding standards for mechanized operations. The Settling Parties propose that the customer service request standard for Verizon be modified. They also propose that the measurable standard for Verizon’s fully electronic data interface (EDI/COBRA) be determined at a future date, and also propose eliminating the standard for Verizon’s Reject/Failed Inquiries.

APPENDIX B
Page 2

The Settling Parties also propose that Pacific's loop qualification standard be modified to reflect their agreement. The Settling Parties also request the addition of language to the “business rules” that will describe the measured interval for Pacific and Verizon, and that will explain that requests for greater than 50 working telephone numbers are excluded for Pacific. In addition, they ask for the addition of language that specifies which interfaces are measured.

The Settling Parties propose adding language to explain that fully electronic processes are measured against system hours, and manual processes are measured against business hours.

The Settling Parties also propose the deletion of language regarding the audit and information submission obligations already met by Pacific and Verizon. The Settling Parties request the deletion of language regarding Verizon's obligation to implement electronic pre-order processes, on the basis that such language defines the duties and rights of parties and, therefore, should not be part of the JPSA. The Settling Parties also ask the Commission to add language that clarifies that Verizon does not support manual engineering query for loop qualifications.

Finally, the Settling Parties propose adding language stating that Service Bureau Provider processing, availability, and response time is not counted against Pacific.

The Settling Parties disagree over a proposal to include "facility availability" information in response to a pre-ordering query, a proposal to measure all loop qualifications queries at parity, a proposal to limit the number of customer service records that can be requested in a single customer service

APPENDIX B
Page 3

record request, a proposal for Verizon to establish a manual loop qualification process, and a proposal to change customer service request measurements for Verizon. The Settling Parties have submitted these disputes for resolution by the Commission. Covad submitted and has since withdrawn a proposal to evaluate Verizon's "Held" and "Denied" sub-measures at parity.

Measurement 2: Average FOC/LSC Notice Interval.

When a CLEC submits an order for local telephone services to the ILECs, Pacific/Verizon respond with a Firm Order Confirmation (FOC) or Local Service Notice (LSC). The FOC/LSC document commits to a due date for service initiation. Measurement 2 captures the time it takes for an ILEC to return a "firm order confirmation" (FOC) or "local service confirmation" (LSC) once it receives a *valid* service request from a CLEC.

The Settling Parties propose examining response times for "valid" service requests alone. The Settling Parties also request adding language to the "reported by" section to reflect Verizon's agreement to report Standalone DSR's as a separate service group type. The Settling Parties propose adding language to the "measurable standard" section to reflect their agreement on the treatment of projects. The Settling parties also propose adding language to the measurable standard that reflects that "Interconnection Trunk Requests – Held and Denied" will be measured for Pacific at parity.

In their July filing, the Settling Parties propose extensive changes to the measurable standards for both Pacific and Verizon. In the November filing, the Settling Parties also propose adding levels of reporting disaggregation for Pacific (i.e. unbundled network element (UNE) Loop – DS3, UNE Loop – OC Level,

APPENDIX B

UNE Dedication Transport – Optical Carrier (OC) Level, Enhanced Extended Links (EEL) – OC Level). They also propose making the measurable standards for Verizon's EEL, Subloop, and Dark Fiber service group types diagnostic.

In addition, The Settling Parties propose making Verizon's measurable standard for "Held and Denied – Interconnection Trunk Requests" a benchmark of 13 days. The Settling Parties request modifying the business rules to reflect their agreement that delays caused by customers are excluded and that loop qualification time for certain products be excluded. They also propose adding "Dark Fiber" to the list of products for which pre-qualification time will be excluded.

The Settling Parties also propose adding language (a) to explain that fully electronic processes should be tracked against system hours; (b) to exclude customer caused delays from the measurement; and (c) to reflect their agreement that days measured will be business days. They also propose adding language to reflect their agreement that the ILEC will perform pre-qualification if pre-qualification has not been completed prior to the submission of the service request by the CLEC. The Settling Parties also seek to delete language regarding projects and interim benchmarks and diagnostic reporting. They also seek to add language that reflects that the Service Bureau Provider processing, availability and response time is not counted against Pacific.

The Settling Parties continue to disagree about proposed new benchmark standards for Verizon's FOCs/LSCs, and submit this dispute for resolution by the Commission.

APPENDIX B
Page 5

Measurement 3: Average Reject Notice Interval.

When a CLEC submits a service request for local telephone services to an ILEC, Pacific/Verizon respond either with an FCO, the subject of Measurement 2, or a notice rejecting the request for service. Measurement 3 reflects the average interval from receipt of a service request to issuance of a rejection notice.

The Settling Parties propose modifying the method of calculating the measurement so that the measurement will reflect certain differences between mechanized and manual rejections. The Settling Parties also seek to update the scope of the measurement by including the high bandwidth line-sharing UNE and standalone directory listings.

Other modifications include adding language (a) to reflect the treatment of projects for Pacific under the "measurable standard" section; (b) to explain time measured for fully electronic processes and manual processes; (c) to exclude customer caused delays; (d) to exclude loop qualification time for certain Pacific products; (e) to reflect the agreement that both Pacific and Verizon will perform pre-qualification if pre-qualification has not been completed prior to the submission of the service request by the CLEC; and (f) to exclude those delays caused by the Service Bureau Provider from being counted against Pacific. The Settling Parties also propose modifying the business rules in their November 2000 filing to state that the loop qualification/facility availability interval is removed from Pacific's overall reject interval for dark fiber.

APPENDIX B
Page 6

The Settling Parties disagree about a proposal that the Commission set a new benchmark for Verizon under this measurement. The Settling Parties have submitted this dispute to the Commission for resolution.

Measurement 4: Percent of Flow Through Orders.

This measurement captures the percentage of electronically received orders that are processed on a flow-through basis, without manual intervention. Measuring flow-through is important because it gauges the efficiency with which Pacific / Verizon are processing CLEC service orders.

The Settling Parties propose treating the measurement as a "diagnostic" standard, and therefore, recommend that the Commission not establish either a benchmark or parity standard. They, however, have proposed re-examining the standard in the course of the next review proceeding. They also recommend excluding orders with syntax, but not content, errors.

There are no “open issues” regarding Measurement 4.

B. Provisioning Measurements.

Measurement 5: Percentage of Orders Jeopardized.

This measurement captures the percentage of orders processed for which Pacific / Verizon notify the CLEC that the order will not be completed by the date committed on the original Firm Order Confirmation (FOC). This measurement bears directly on the ability of CLECs to communicate accurate information to their customers.

The Settling Parties propose reporting the data captured by this measurement by Service Group Type only, and not by interface type or type of jeopardy. Thus, they request that the Commission adopt new language defining

APPENDIX B
Page 7

the measurable standard, which will reflect their agreement on this issue. They also propose adding levels of disaggregation for Pacific's reports. For Verizon, the Settling Parties also propose including language that will clarify the "retail comparison" for local number portability (LNP) by adding the words "Total Business and Residence, Non Dispatched." They also propose amending the business rules to add language that will explain that raw data will include jeopardy codes, that UNE subloop will be tracked diagnostically, and that dark fiber will be tracked diagnostically until the next periodic review. The Settling Parties also ask for the addition of language to clarify that the measurement does not capture "missed commitments."

The Settling Parties have been unable to agree about a proposal that Verizon and Pacific report results for conditioned and non-conditioned loops on disaggregated bases for digital subscriber line (xDSL) loops. The Settling Parties have submitted this dispute for resolution to the Commission.

Measurement 6: Average Jeopardy Notice Interval.

If Pacific detects that it probably will not meet the due date for service installation specified in its Firm Order Confirmation (FOC), it issues a notice to the CLEC indicating the order is in jeopardy of missing the due date. Measurement 6 captures the average time between the completion date an ILEC states in its FOC and the date and time the ILEC issues either (a) a notice to the CLEC that the order is in jeopardy of missing the due date; or (b) a notice indicating that the due date has already been missed.

The Settling Parties have proposed adding language to clarify the method of calculation of this measurement as well as language which would limit reporting to service group types, instead of also reporting by interface type or type of jeopardy. The Settling Parties also propose modifying the description of “Assignment” jeopardy and “Installation” jeopardy under the “Method of Calculation” section. The Settling Parties also propose a benchmark for Pacific. The Settling Parties also request that Verizon track data for four months, at the end of which benchmarks will be set on the basis of the four months review.

The Settling Parties propose adding additional levels of reporting disaggregation for both Pacific and Verizon under the "Measurable Standards" section. They also propose making Verizon's EEL, Subloop and Dark Fiber measurements diagnostic in nature. The Settling Parties also propose to delete unnecessary language. They also propose adding business rules regarding the method by which orders classified as in jeopardy are tracked. Furthermore, they propose a description of how a jeopardy is treated on the due date for Verizon.

The Settling Parties continue to disagree about the proposal that Pacific and Verizon report results for conditioned and non-conditioned loops on desegregated bases for xDSL loops. The Settling Parties have submitted this dispute to the Commission for resolution.

Measurement 7: Average Completed Interval.

Measurement 7 examines the average number of business days it takes an ILEC to complete a valid service request, as reflected by the number of business days between the date requested and the date of completion reflected in the service order system.

The Settling Parties propose that the Commission adopt language that (a) delineates the service group types that should be reported; (b) excludes orders

APPENDIX B
Page 9

that have an interval different from the offered interval; (c) addresses the treatment of projects; (d) mandates a diagnostic tracking of dark fiber and UNE subloops (except X-DSL for Pacific) for both ILECs and for Verizon EELs; and (e) with regard to UNE loop services, excludes feature only orders from the retail analog.

The Settling Parties propose further disaggregation of Pacific’s reporting as well as adding sub-measures for Pacific’s xDSL, UNE Loops, and Line Sharing reports. They also propose clarifying Verizon’s retail comparison for LNP to include the words, “Total Business and Residence, Non-Dispatched.”

The Settling Parties also propose modifying language to reflect what they submit is the appropriate analog for DSL services. For Pacific, the Settling Parties also propose adding a business rule regarding the relevance of “Completion Date” to “Acceptance Testing.” They also remove language from the “Notes” section which is no longer relevant.

The Settling Parties continue to disagree about a proposal about the definition of a "completion date" under circumstances when an "acceptance test" has been requested. Pacific has accepted a modified version of Covad's recommendation on this point, but Verizon continues to reject it. The Settling Parties submit this issue as it applies to Verizon for resolution by the Commission. Covad has also proposed that Verizon report results for conditioned and non-conditioned loops on disaggregated bases for xDSL loops.

APPENDIX B
Page 10

Measurement 8: Percent Completed Within Standard Interval.

This measurement examines the number of received, valid orders completed within a standard interval. This measurement complements information provided by measuring the Average Completed Interval and suggests the extent to which service completion times vary from an expected timeframe.

The Settling Parties propose adjusting the JPSA's language to reflect their consensus on the service group types they say should be reported. They also propose adding additional levels of disaggregation to Pacific's reports. They request that the Commission change the language of the business rules and exclude orders that have an interval different from the standard interval.

In the revised JPSA, they propose adding language that would require diagnostic tracking of UNE subloops and dark fiber for Pacific. The Settling Parties also seek to add language that will exclude "feature only" orders from the retail analog for UNE loop services. The Settling Parties propose deleting language regarding projects as well as modifying language to reflect their consensus regarding the appropriate analog for DSL services. The Settling Parties also propose modifying the "business rules" by adding a new rule for Pacific Bell which explains the relevance of "Completion Date" to "Acceptance Testing."

In their comments, Covad and NorthPoint propose a completed interval benchmark of 95% within 7 days for non-conditioned loops and 11 days for conditioned loops for Verizon's xDSL UNE loops and line sharing UNE. They also propose that the Commission establish for Pacific a completed interval

APPENDIX B
Page 11

benchmark of 95% within 5 days for non-conditioned loops and 10 days for conditioned xDSL UNE loops. Covad recommends that Pacific and Verizon report results for conditioned and non-conditioned loops on disaggregated bases for xDSL loops. Covad also seeks a modification of the definition of "completion date" under circumstances where an "acceptance test" has been administered.

Pacific has agreed to a modified version of Covad's original proposal, but Verizon continues to reject the proposal. Covad’s issues with Verizon are before the Commission for resolution.

Finally, Covad proposes establishing standard intervals by service group types for Verizon's UNE services that would result in the inclusion of UNE services within this measurement. The Settling Parties do not agree on these proposals and submit them to the Commission for resolution.

Measurement 9: Coordinated Customer Conversion.

Coordinated orders require Pacific/Verizon to disconnect a customer's service. As such, the importance of Pacific/Verizon's completion of a coordinated conversion service order at the committed date and time lies in the fact that a CLEC needs to be prepared to immediately begin migrating a customer's service in order to prevent the customer from going without service. This measurement tracks the percentage of coordinated "cutovers" completed by Pacific by the committed time. The measurement also captures the percentage of coordinated orders completed by Verizon before or at the committed time.

The Settling Parties propose modifying the description of the measurement to specify that the measurement captures "cutovers" by Pacific. The Settling Parties have proposed refining the method of calculation for Verizon

APPENDIX B
Page 12

as well as the reporting structure for Pacific. The Settling Parties also seek to add language that clarifies the Pacific measure as well as defines certain terms under the Verizon measure. The Settling Parties propose the introduction of language to define "cutovers." The Settling Parties request the substitution of the term "local number portability" for "permanent number portability", the former of which is the more up-to-date technical expression.

Following the February 13, 2001 agreement between Verizon and the participating CLECs, there is no longer an open issue with respect to Measurement 9.

Measurement 9A: Frame Due Time (FDT) Conversions as Percentage on Time (Pacific Bell Only).

The Settling Parties have proposed an additional coordinated cutover measurement that examines the percentage of the number of frame due time (FDT) cutovers completed by Pacific within the initial time commitment. The Settling Parties propose calculating this measurement as the factor of 100 and the quotient of the number of frame due time cutovers completed by the committed time and the count of frame due time cutovers scheduled within a reporting period, which they suggest should be one month.

The Settling Parties propose that reports be structured to reflect results by individual CLECs, CLECs in the aggregate, Pacific, and Pacific affiliates. They propose that reports address basic loops with and without local number portability, and standalone local number portability. They seek to report results on a statewide basis. The Settling Parties request a benchmark of 95% in one hour. They also propose two business rules which would exclude CLEC caused misses and which limit the scope of the measurement to CLEC requested FDT orders. The Settling Parties also define "cutovers" to include initial and subsequent attempts to complete a cutover. The measurement will cover up to 19 loops or up to 99 telephone numbers on standalone local number portability.

There are no open issues regarding Measurement 9A.

Measurement 10: LNP Network Provisioning.

This measurement calculates the success rate for local number portability (LNP) network provisioning. LNP is critical to the successful development of competition in the local telephone markets. When Pacific/Verizon fail to provide LNP, customers switching to another local carrier face the possibility of interrupted service, and therefore, will have an incentive to continue purchasing services from their current providers.

The Settling Parties have proposed updating the term "permanent number portability" to reflect current usage, which is "local number portability." The Settling Parties have also sought the addition of language which would set benchmark measurements for Pacific and Verizon. Furthermore, the Settling Parties request the modification of language (a) concerning the tracking of provisioning failures; (b) limiting the broadcast exclusions to Pacific; (c) excluding large porting activities for Pacific; and (d) deleting a note regarding the implementation and timing of Verizon's reporting requirement because it is no longer relevant.

There are no open issues regarding Measurement 10.

APPENDIX B
Page 14

Measurement 11: Percent of Due Dates Missed.

This measurement examines the percentages of CLEC orders that are not completed by the due date listed on the firm order confirmation. It measures both the accuracy of the information transmitted on the firm order confirmation and the timeliness with which Pacific/Verizon complete CLEC service orders.

The Settling Parties propose adding language to reflect their agreement about the service group types that should be reported. They also request the addition of language that reflects their agreement on the exclusion of "feature only" orders from Pacific's retail analog for the UNE loop. The Settling Parties also propose refining the levels of disaggregation of Pacific's reports. They also propose to clarify Verizon's retail comparison for LNP by adding the words, "Total Business and Residence, Non-Dispatched." The Settling Parties propose the addition of language that treats dark fiber as a diagnostic measurement.

The Settling Parties also propose adding language (a) about the "record only" ILEC official orders; (b) that would require ILECs to provide disaggregation by missed appointment when requested to do so in a raw data request; (c) concerning a business rule that would clarify the link between "Completion Date" and "Acceptance Testing" for Pacific; and (d) which explains why the retail comparison for Integrated Services Digital Network (IDSL) capable UNE loops is ISDN. Finally, the Settling Parties propose deleting language regarding the analog because it is unnecessary.

The Settling Parties disagree about a proposed recommendation that the results for conditioned and non-conditioned loops be reported on disaggregated bases for the xDSL loops of both Pacific and Verizon and have submitted this as an open issue.

Measurement 12: Percent Due Dates Missed Due to Lack of Facilities.

This measurement is a subset of Measurement 11. It calculates the percentage of due dates that were missed because of a lack of facilities.

The Settling Parties have proposed the addition of language to reflect their agreement about the reporting of service group types. They propose the addition of language that would reflect their consensus regarding the exclusion of "feature only" orders from the retail analog for UNE loop services.

The Settling Parties also request the modification of language regarding the appropriate analog for DSL services. The Settling Parties also propose adding levels of disaggregation to Pacific's reports.

The Settling Parties disagree about a recommendation that Pacific include UNE Subloop disaggregation for this measure. This open issue is now, therefore, before the Commission.

Measurement 13: Delay Order Interval to Completion Date (For Lack of Facilities).

This measurement examines the average number of calendar days that elapse from the due date to completion date due to lack of ILEC facilities.

The Settling Parties propose (a) adding language on the measurement standards for service group types and their agreement regarding the exclusion of feature only orders from the retail analog for UNE loop services; (b) modifying language regarding the appropriate analog for DSL services; (c) adding several new levels of disaggregation to Pacific's reports; and (d) clarifying under the

APPENDIX B
Page 16

"Measurable Standard" that Verizon's retail comparison for the UNE Port is "CentraNet-Simple."

The Settling Parties disagree about a recommendation that Pacific include UNE Subloop disaggregation for this measure. This open issue is now, therefore, before the Commission.

Measurement 14: Held Order Interval.

This measurement examines the average time service orders are left incomplete because of ILEC-related reasons, including lack of facilities. It looks back from the completion date to determine how long the request was left pending. The Settling Parties propose adding language (a) about the measurable standards for service group types; (b) that would clarify that Verizon's retail comparison for UNE Port is "CentraNet-Simple"; to Verizon's retail comparison for LNP; (c) excluding "feature only" orders from the retail analog for UNE loop services. The Settling Parties also propose modifying language regarding the appropriate analog for DSL services, and adding language that would reflect their agreement that the UNE subloop and dark fiber be tracked as diagnostic measurements. The Settling Parties also propose adding business rules for Pacific which clarify the connection between "Completion Date" and "Acceptance Testing." The Settling Parties also propose that the ILECs disaggregate raw data by missed appointment codes when requested to do so. There are no open issues for Measurement 14.

Measurement 15: Provisioning Trouble Reports.

This measurement captures the number of trouble reports received from a customer (or indirectly through the CLEC the customer has migrated to)

APPENDIX B
Page 17

that occur from the time that a CLEC places a service order request with Pacific / Verizon until the time the service order is completed. It allows the Commission to compare Pacific / Verizon's processing of competitor's service orders to the manner in which Pacific / Verizon handle service orders for their own retail customers. The Settling Parties propose modifying the method of calculation by creating distinct formulas for parity and benchmark sub-measurements. The Settling Parties also request modifications to language regarding the reporting of service group types, and about the measurable standards for both ILECs' service group types. The Settling Parties propose language to clarify the benchmarks for LNP for Pacific and Verizon.

The Settling Parties also propose adding language that will indicate the availability of additional data if, and when, a CLEC requests it. They propose deleting language regarding Verizon programming and reporting obligations because the language is inappropriate for the JPSA, and deleting language about the development of measurements, because the language is no longer relevant.

The Settling Parties cannot agree about recommendations that (a) Pacific / Verizon report new services troubles prior to the completion of service orders; (b) parity with Verizon serve as a measurable standard for the local number portability sub-measure; (c) results for Verizon / Pacific's conditioned and non-conditioned loops be reported on disaggregated bases for xDSL loops and line shared loops; and (d) a parity comparison with ASI for Pacific's xDSL sub-measures serve as the measurable standard. The Settling Parties submit these disputes for resolution by the Commission.

APPENDIX B
Page 18

Measurement 15A: Average Time to Restore Provisioning Troubles.

This is a new measurement proposed by the Settling Parties, which would examine how long it takes ILECs to resolve problems during the provisioning process. Measurement 15 examines the *frequency* of provisioning troubles. Measurement 15A calculates the average *duration* of trouble by dividing the duration of all provisioning troubles from the time the trouble began by the number of reports of provisioning trouble.

The Settling Parties propose reporting this measurement on a monthly basis for individual CLECs, CLECs in the aggregate, individual ILECs, and all ILEC affiliates. The measurable standard for Pacific is both parity (for Resale and UNE Loop) and a benchmark (for LNP), and it is a retail comparison for Verizon. The Settling Parties also propose that the business rules exclude CPE and IEC/CLEC caused troubles, subsequent reports, message reports, and reports generated by ILEC employees, and that raw data be disaggregated by maintenance disposition codes, when so requested by a CLEC.

The Settling Parties continue to disagree over a proposal that a parity comparison with Pacific's affiliate, ASI, serve as the measurable standard for xDSL and line shared loops. They also disagree over the recommendation that results for Verizon's and Pacific's conditioned and non-conditioned loops should be reported on disaggregated bases for xDSL loops and line shared loops. The Settling Parties have submitted this dispute to the Commission for resolution.

APPENDIX B
Page 19

Measurement 16: Percent Troubles in 30 Days for New Orders (Specials).

The Settling Parties propose revising Measurement 16 to make it strictly applicable to special services. Measurement 16 used to apply to all services for Pacific and designed services for Verizon. Measurement 17 used to apply to non-designed services for Verizon. The Settling Parties suggest making Measurement 16 the gauge for special services for both ILECs and Measurement 17 the gauge for non-special services for both ILECs.

The Settling Parties propose adding language to Measurement 16 that (a) would clarify the types of orders included in this measure; (b) the method of calculation captures only special services orders; (c) would include xDSL, UNE Loops, IDSL UNE Loops, and Line Sharing under this measure for Verizon; and (d) would address service group types. The Settling Parties propose adding several new levels of disaggregation to Pacific's reports.

The settling parties also seek to add language to the "business rules" that would reflect their agreement on necessary adjustments that Pacific would make when no orders are processed in a given month. Other changes include adding language that explains the connection between "completion date" and "acceptance testing," and adding language that would clarify that additional data from the ILECs would be made available upon request. They also seek to delete language that would pertain to non-special services, and add language that would emphasize that tracking results for UNE subloops and dark fiber would be done solely for diagnostic purposes until the next review period.

Initially the Settling Parties indicated that they could not agree about a recommendation that Verizon include xDSL when measuring percentage of

APPENDIX B
Page 20

troubles in 30 days. They submitted their dispute to the Commission for resolution. As evidenced by their proposal in their November 6, 2000 "Submission," which would include xDSL under this measurement for Verizon, the Parties have reached an agreement on this issue. The Commission will treat this as a "closed" issue. Therefore, there are no open issues regarding Measurement 16.

Measurement 17: Percent Troubles in 7 (GTE) or 10 days (Pacific) for Non-Special Orders.

The Settling Parties suggest adjusting the scope of Measurement 17 to make it the gauge for troubles with non-special services of both ILECs. Previously it applied only to non-designed services of Verizon. They propose adding language that clarifies the types of orders included in this measurement, and the method of calculation by the ILEC. The Settling Parties also seek to add language to the measurable standard that would reflect their agreement about service group types.

They propose changing the business rules to reflect their agreement on the necessary adjustments that Pacific should make when it processes no orders in a given month. The Settling Parties also seek to add language to clarify that additional data is available from the ILECs on request, as well as language that FDT and TBCC should be tracked diagnostically for Pacific. They also propose adding language that results in UNE subloops being tracked diagnostically until the next review period. The Settling Parties also propose (a) making the retail comparison for UNE Platform – Basic port and Loop for Pacific to "Business (disp / non-disp); (b) excluding xDSL, UNE Loops, IDSL UNE Loops, and Line

APPENDIX B
Page 21

Sharing from Verizon's reports under this measurement; (c) changing Verizon's benchmark for LNP to a parity measurement; and (d) adding a business rule for Pacific that explains the conceptual connection between "Completion Date" and "Acceptance Testing." There are no open issues regarding Measurement 17.

Measurement 18: Completion Notice Interval.

This measurement captures the percent of completion notices returned within the time specified in the measurable standard.

The Settling Parties propose revising the language of the measurement so that the measurement should now be reported as a percentage figure, not an average. The Settling Parties also propose reporting this measurement for all interfaces for both ILECs and modifying the language of the measurement standard to report the measurement as a percentage instead of an average figure. They also offer a new standard for Pacific for electronic orders that fall out for manual processing. The Settling Parties request the addition of language to explain that system hours be used to measure fully electronic submeasures. The Settling Parties propose deleting language regarding interim benchmarks and Verizon's programming and reporting obligations as inappropriate for the JPSA. The Settling Parties also propose modifying the benchmark standards for Verizon. They also propose adding business rules to clarify Verizon's CN reporting process, and re-writing the notes to clarify that retail disconnects are included under this measurement. Finally, the Settling Parties propose adding language that this measurement does not pertain to disconnect orders placed by the ILEC.

APPENDIX B
Page 22

The Settling Parties submitted a proposal for resolution that would have established a benchmark for Verizon's fully electronic submeasures. After their submission, Verizon and the CLECs have indicated that they now agree to the following benchmarks for Verizon:

95% within 1 hour for fully electronic, such as EDI;
95% within 12 hours for other electronic, such as WISE;
90% in 24 hours for other manual processes.

There are no open issues regarding Measurement 18.

C. Maintenance Measurements.

Measurement 19: Customer Trouble Report Rate.

This measurement calculates the number of network customer trouble reports in a calendar month, as a percentage of the total number of access lines/circuits/UNEs in service at the end of the prior reporting period. The measurement allows the Commission and the parties to compare the quality of facilities and services provided to CLECs and their customers with those provided to Pacific/Verizon customers. The Commission can thereby ensure that Pacific/Verizon is providing CLECs with services and facilities in a non-discriminatory fashion.

The Settling Parties propose (a) modifying the language of the measurement to reflect the current terminology for number portability; (b) having the measurable standard reflect their agreement regarding service group types; and (c) expanding the levels of disaggregation of Pacific’s reports. Furthermore, the Settling Parties request that the business rules reflect that Verizon's results exclude provisioning trouble reports. The Settling Parties also propose that both ILECs include Test-OK (TOK) and Found-OK (FOK) reports under this measurement.

The Settling Parties also propose (a) adding language that will clarify that additional data from the ILEC is available upon request; (b) deleting language regarding the appropriate analog for DSL services and the development of the measure; and (c) adding language which classifies results for UNE subloops and dark fiber as diagnostic measurements.

There are no open issues under Measurement 19.

Measurement 20: Percent Customer Trouble Not Resolved Within Estimated Time.

This measurement captures the percentage of troubles reported which are not resolved within the time committed to by Pacific/Verizon. The measurement compares the timeliness with which Pacific/Verizon respond to CLEC customer troubles with the timeliness with which Pacific/Verizon respond to troubles reported by Pacific/Verizon customers. It thus enables the Commission and the parties to evaluate the extent to which CLEC customer troubles are resolved in a timely, non-discriminatory fashion.

The Settling Parties propose (a) modifying the language of the measurement to reflect the current terminology for number portability; (b) having the measurable standard reflect their agreement regarding service group types; and (c) adding several new levels of disaggregation to Pacific's reports under this measurement. Furthermore, the Settling Parties recommend that the business rules reflect that Verizon's results exclude provisioning trouble reports. The Settling Parties also request that both ILECs include “Test-OK” and “Found-OK” reports under this measurement.

The Settling Parties also propose (a) adding language that clarifies that additional data from the ILEC is available upon request by a CLEC; (b) deleting language regarding the appropriate analog for DSL services and the development of the measure; and (c) adding language which classifies results for UNE subloops and dark fiber as diagnostic measurements.

There are no open issues under Measurement 20.

Measurement 21: Average Time to Restore.

This measurement calculates average duration of customer trouble reports, and thus complements Measurement 20 above, which measures the percent of trouble reports resolved in a committed timeframe. The measurement compares the timeliness with which Pacific/Verizon respond to CLEC customer troubles with the timeliness with which Pacific/Verizon respond to troubles reported by their own retail customers. It thus enables the Commission and the parties to evaluate the extent to which CLEC customer troubles are resolved in a timely, non-discriminatory fashion.

The Settling Parties propose (a) modifying the language of the measurement to reflect the current terminology for number portability; (b) having the measurable standard reflect their agreement regarding service group types; and (c) adding several new levels of reporting for Pacific. Furthermore, the Settling Parties request that the business rules reflect that Verizon's results exclude provisioning trouble reports. The Settling Parties also propose that both ILECs include “Test-OK” and “Found-OK” reports under this measurement.

The Settling Parties also propose (a) adding language that will clarify that additional data from the ILEC is available upon request; (b) deleting language regarding the appropriate analog for DSL services and the development of the measure; and (c) adding language which classifies results for UNE subloops and dark fiber as diagnostic measurements. The Settling Parties also seek to change Verizon’s LNP retail benchmark to a parity standard.

The are no open issues under Measurement 21.

Measurement 22: POTS Out of Service Less Than 24 Hours.

This measurement captures the percentage of Plain Old Telephone Service (POTS) out-of-service trouble reports that are resolved within 24 hours of the report. This measurement enables the Commission and the parties to compare the timeliness with which CLEC POTS troubles are resolved with the timeliness with which Pacific/Verizon resolve POTS troubles for their own customers.

The Settling Parties propose adding language to reflect their agreement regarding service group types, as well as language to reflect their agreement that Pacific's UNE subloops be tracked diagnostically by UNE loop type. Results will also include TOK and FOK reports for both ILECs.

There are no open issues under this Measurement 22.

APPENDIX B

Measurement 23: Frequency of Repeat Trouble in 30-Day Period.

This measurement evaluates whether troubles are chronic in nature by capturing the percentage of repeat troubles reported within 30 days of a previous report. The measurement compares the effectiveness with which Pacific / Verizon resolve troubles reported by Pacific / Verizon customers with their effectiveness in resolving troubles reported by CLECs and their customers. It thus enables the Commission and the parties to evaluate whether Pacific / Verizon are resolving CLEC customer troubles in an effective, non-discriminatory fashion.

The Settling Parties propose (a) updating language to reflect the current industry term for number portability; (b) adding language to reflect their agreement about service group types; (c) adding language to clarify that additional data is available from the ILEC upon request in conjunction with a CLEC's request for raw data; (d) deleting language regarding the appropriate analog for DSL services; and (e) expanding the disaggregation of Pacific's reports.

There are no open issues under Measurement 23.

D. Network Performance Measurements.

Measurement 24: Percent Blocking on Common Trunks.

This measurement evaluates the percentage of common and shared trunk groups with blockage in excess of 2%.

The Settling Parties propose (a) modifying language to reflect their agreement to report by total trunk group on a statewide basis; (b) adding language to reflect their agreement on reporting requirements that will provide detailed information for all trunk groups not meeting the 2% level; and (c) deleting Notes section of the measurement as no longer relevant.

The are no open issues under Measurement 24.

Measurement 25: Percent Blocking on Interconnection Trunks.

This measurement captures the percentage of dedicated interconnection trunks which experience blockage in excess of 2%. Quality network transmission is essential to a CLEC's success in a local telephone market. This measurement allows the Commission to ensure that the networks operate at a level sufficient to support a competitive environment and that Pacific / Verizon allocate trunk capacity on a non-discriminatory basis.

The Settling Parties have proposed (a) modifying language to reflect their agreement that total trunk groups be reported by individual CLEC on a statewide basis; (b) adding language that reflects their agreement to exclude failures caused by a CLEC that fails to complete growth trunk provisioning by scheduled due date; (c) changing language in the business rules section to explain when the measure applies and what it excludes; and (d) deleting language from the notes as no longer relevant.

There are no open issues under Measurement 25.

Measurement 26: NXX Loaded by LERG Effective Date.

This measurement calculates the number of telephone number prefixes (NXXs) loaded and tested by the Local Exchange Routing Guide Effective Date (LERG). LERG is an independent database that serves the telecommunications industry. It provides standard time intervals for the loading and testing of NXXs. Pacific's/Verizon's loading of a competitor's NXX is necessary if Pacific/Verizon customers are to be able to call the competitor's customers with that NXX. This measurement allows the Commission and the parties to compare the timeliness with which Pacific/Verizon load and test CLEC NXXs with the timeliness with which Pacific/Verizon load their own NXXs. It likewise allows the Commission to evaluate the efficiency with which Pacific/Verizon are accomplishing this important task.

The Settling Parties propose modifying the language to reflect their agreement to exclude NXX codes that cannot be completely tested because the CLEC has not provided accurate test numbers or the CLEC facilities have not been installed and adding language that would include additions and deletions to NXX codes to the measurement.

There are no open issues under Measurement 26.

Measurement 27: Network Outage Notification.

This measurement captures the average interval between a network outage and notification of a CLEC by Pacific/Verizon of the outage. This measurement compares the efficiency with which Pacific/Verizon notify their own departments of an outage with the efficiency with which Pacific/Verizon notify CLECs of an outage of the same type, and thereby allows the Commission and the parties to ensure that CLECs are notified of outages in a prompt and non-discriminatory fashion.

The Settling Parties request the deletion of this measurement in favor of Pacific/Verizon using email notification simultaneously to their own departments and wholesale customers.

E. Billing Measurements.

Measurement 28: Usage Timeliness.

This measurement captures the average time it takes Pacific/Verizon to report usage by a CLEC customer. The measurement is calculated as the time elapsed between the time Pacific/Verizon record of usage by a CLEC customer and when the data is transmitted to the CLEC in compliant form. Timely transmission of usage data is necessary for CLECs to be able to bill their customers. This measurement allows the Commission and the parties to ensure that Pacific/Verizon are transmitting CLEC customers usage data in a non-discriminatory, timely fashion.

The Settling Parties propose modifying the language of the measurement to make the measurable standard a parity standard for most reported services. Under the “Measurable Standard” section, the Settling Parties propose that Verizon document separate sub-measures of the UNE Platform-Local and UNE Platform- Access. The Settling Parties also propose adding language to the “notes” section which will clarify Verizon’s process for local/toll billing documentation.

The Settling Parties initially failed to agree about a proposal that Verizon establish a new level of disaggregation for UNE-Access.

There are no open issues for resolution under Measurement 28.

APPENDIX B
Page 30

Measurement 29: Accuracy of Usage Feed.

This measurement captures the completeness of content, accuracy of information, and correctness of formatting of usage records transmitted by Pacific/Verizon to CLECs. Accuracy of usage records enables CLECs to promptly and correctly bill their customers, an important element in the CLECs' ability to provide quality competitive service. This measurement thus enables the Commission and the parties to ensure that Pacific's/Verizon's recording and transmittal of CLEC usage data meet a high standard of quality sufficient to support a competitive local telephone market.

In our earlier decision (D.99-08-020), we directed the parties to establish criteria for the measurement and postpone setting a benchmark until then. The Settling Parties proposed that (a) the measurement be reported as a percentage of all usage records received and processed and that the measurement be reported on a monthly basis; (b) the Commission defer setting a measurable standard until the next review period or until three months of data are collected, whichever comes first; and (c) we add several new business rules.

There are no open issues for resolution under Measurement 29.

Measurement 30: Wholesale Bill Timeliness.

This measurement captures the number of days between the close of the billing cycle and the date Pacific/Verizon transmit the bill to the CLEC. This measurement enables the Commission and the parties to ensure that Pacific's/Verizon's wholesale billing of CLEC usage meets a high standard of quality sufficient to support a competitive local telephone market.

APPENDIX B
Page 31

The Settling Parties request modifying the language of the measurement in order to clarify that the measurement will examine calendar days, not business days, and adding language that reflects their agreement that Verizon will report UNE and Resale as a combined result.

The Settling Parties disagree about a proposal that sub-measures be established for Pacific's / Verizon's paper, magnetic, CD-ROM and Custom Bill diskette bills. They have submitted this issue to the Commission for resolution.

Measurement 31: Usage Completeness.

This measure captures the percentage of usage charges which appear on the correct bill. Timely, complete billing of usage enables CLECs to promptly and correctly bill their customers and collect accurate internal financial data, important elements in the CLECs' ability to provide competitive service. This measurement enables the Commission and the parties to ensure that Pacific's / Verizon's transmittal of usable bills is sufficiently complete and timely to support a competitive local telephone market.

The Settling Parties propose adding language to adjust the time period for capturing data for Pacific and adding language to reflect that Verizon will report UNE and Resale as a combined result.

There are no open issues under Measurement 31.

Measurement 32: Recurring Charge Completeness.

This measurement captures the percentage of recurring charges which appear on the correct bill. Timely, complete billing of recurring charges enables CLECs to promptly and correctly bill their customers and collect accurate internal financial data, important elements in the CLECs' ability to provide competitive service. This measurement enables the Commission and the parties to ensure that Pacific's/Verizon's transmittal of recurring charge bills is sufficiently complete and timely to support a competitive local telephone market.

The Settling Parties propose (a) adding language indicating that Verizon will calculate this measurement using dollar amounts; (b) modifying the language of Verizon's measurable standard; (c) adding language that reflects their agreement to exclude mandated billing changes; and (d) adding language to reflect their agreement that the measurement will be retired for Pacific 60 days after it begins reporting the proposed new measurement, Measurement 35.

There are no open issues under this Measurement 32.

Measurement 33: Non-Recurring Charge Completeness.

This measurement captures the percentage of non-recurring charges which appear on the correct bill.

The Settling Parties propose (a) adding language indicating that Verizon will calculate this measurement using dollar amounts; (b) modifying the language of Verizon's measurable standard; (c) adding language that reflects their agreement to exclude mandated billing changes; and (d) adding language to reflect their agreement that the measurement will be retired for Pacific 60 days after it begins reporting the proposed new measurement, Measurement 35.

There are no open issues under Measurement 33.

Measurement 34: Bill Accuracy.

This measurement evaluates the accuracy of Pacific/Verizon billing of CLEC usage by calculating the percentage of monies billed without corrections. Accurate billing by Pacific/Verizon enables CLECs to promptly and correctly bill

APPENDIX B
Page 33

their customers, an important element in the CLECs' ability to provide competitive service.

The Settling Parties propose adding language that reflects their agreement to exclude mandated billing changes and language that reflects their agreement that Verizon will report UNE and Resale as a combined result.

There are no open issues under Measurement 34.

Measurement 35: Duplicate Billing

The Settling Parties propose replacing this measurement, which captures the number of former Pacific customers who receive erroneous bills after conversion to a CLEC service, with a new measurement that captures the timeliness of billing completion notices. The Settling Parties propose that after Pacific implements a billing completion notice process, it will cease reporting under Measurement 32 and 33, sixty days after it commences reporting under the new Measurement 35.

There are no open issues under this measurement.

Measurement 36: Accuracy of Mechanized Bill Feed.

This measurement evaluates the accuracy of mechanized bill feeds. In our earlier decision (D.99-08-020), we directed the parties to develop a set of criteria for this measurement.

The Settling Parties now propose that the measurement be reported by individual CLEC and CLECs in the aggregate and that data be collected and appropriate benchmarks discussed at the next review or after three months of data has been collected, whichever comes first.

There are no open issues under Measurement 36.

F. Database Updates Measurements.

Measurement 37: Average Database Update Interval.

This measurement captures the interval between the time when CLECs submit information updates, to the time when Pacific/Verizon pass the updated customer information to the directory assistance/directory listing databases.

The Settling Parties propose that (a) Pacific track LIDB service order generated updates; (b) language is added that creates a benchmark for direct gateway updates; (c) language is added to specify that the measurement reflects calendar days; and (d) language is updated to reflect Verizon's compliance with certification.

There is an open issue between the CLECs and Verizon about whether Verizon should be required to include LIDB under this measure.

Measurement 38: Percent Database Accuracy.

This measurement calculates the percentage of Emergency 9-1-1 and Directory Assistance/Directory Listings updates completed without error.

The Settling Parties propose adding language that reflects Pacific's agreement to track LIDB service order generated updates and deleting language to reflect Verizon's compliance with the independent audit ordered in D.99-08-020.

The Settling Parties have been unable to agree about a proposal that Verizon add LIDB and MSAG to the list of databases it will measure. Nor have they been able to agree that the measurement be eliminated because it is at parity by design. The Settling Parties have submitted these issues to the Commission for resolution.

Measurement 39: E911/911 MS Database Update.

This measurement examines the efficiency with which Pacific/Verizon update Emergency 9-1-1 databases.

The Settling Parties propose adding language to clarify that service order generated updates are for Pacific only. They also propose that both Pacific and Verizon track direct gateway updates. The Settling Parties seek to clarify the Emergency 9-1-1 processing intervals.

There are no open issues under Measurement 39.

G. Collocation Measurements.

Measurement 40: Time to Respond to a Collocation Request.

This measurement captures the average time Pacific/Verizon take to respond to a CLEC request for collocation. The measurement calculates response time to two kinds of requests, namely, space availability and price/schedule quote requests.

The Settling Parties propose (a) adding language that reflects separate standards for Space Availability and Price/Schedule Quote requests; (b) adding language to specify that the measurement be reported in terms of calendar days; (c) adding language to reflect their agreement on the treatment of revised applications; (c) changing language to identify the impact of collocation request changes on processing intervals associated with power, heating, ventilation, and air conditioning (HVAC), and major building modifications; and (d) adding language to reflect the effect of large orders on Pacific's cageless collocation request processing; and (e) deleting the word “valid” before the words "published ILEC guidelines" in the section. The Settling Parties also propose to treat changes to a collocation application filed with Verizon after a 15 calendar day period as a new application for measurement purposes.

The CLECs and Verizon disagree over a proposal that would adjust the response intervals when ILECs receive ten or more applications within a ten-day period from an individual CLEC. The CLECs and Verizon have submitted this issue to the Commission for resolution.

Measurement 41: Time to Provide a Collocation Arrangement.

This measurement captures the average time it takes Pacific/Verizon to complete or build a collocation arrangement, both for (a) a new arrangement and (b) augmentation of an existing arrangement.

The Settling Parties propose (a) adding language to report the measurement in terms of calendar days; (b) documenting a separate sub-measure for cageless collocation under the "report by" section; (c) adding language that reflects their agreement to exclude requested due dates greater than standard interval; (d) adding language that reflects their agreement on the effect of large orders on Pacific's cageless collocation construction intervals; (e) adding a business rule which will explain the effect of CLEC delays on Pacific's reporting of collocation construction intervals; and (f) establishing new sub-measures for cageless collocation at Pacific premises.

The Settling Parties do not agree about a proposal to reduce the actual installation interval when a CLEC changes the collocation request and that change results in an interval longer than the committed installation interval.

Pacific has agreed to a slightly modified version of the original proposal. Nor do they agree about a proposal to redefine the levels of disaggregation for Verizon collocation requests. The Settling Parties do not agree about a proposal to establish new benchmarks for Verizon's provisioning intervals. Finally, they do not agree about a proposal to establish new sub-measures for cageless collocation at Verizon premises. The Settling Parties have submitted these issues to the Commission for resolution.

H. Interface Measurements.

Measurement 42: Percent of Time Interface is Available.

This measurement evaluates the accessibility of Pacific's/Verizon's OSS systems during the time in which they are scheduled to be available. The Settling Parties propose rewording the measurement to calculate the impact on "interfaces" instead of "systems" and adding language that reflects their agreement that ILECs report affiliate data. They also propose that Verizon report data on a nationwide basis.

There are no open issues under Measurement 42.

Measurement 43: Average Notification of Interface Outages.

This measurement calculates the average time it takes for Pacific/Verizon to notify the CLECs that Pacific's/Verizon's OSS interface is experiencing an outage.

The Settling Parties propose eliminating this measurement altogether. They propose establishing a "parity by design" process which would involve e-mailing notice of outages simultaneously to retail and wholesale customers.

There are no open issues under Measurement 43.

Measurement 44: Center Responsiveness.

This measurement captures the average time it takes for Pacific's / Verizon's ordering and repair centers to respond to a CLEC call.

The Settling Parties propose (a) adding language that reflects their agreement that Pacific report by provisioning center; (b) modifying Verizon's benchmark and adding language to reflect Pacific's agreement to report for the provisioning center as well as Pacific's agreement to a benchmark for this new sub-measure; (c) adding language to reflect that Verizon will report data on a nationwide basis; and (d) adding language to the “notes” section describing Verizon’s two repair centers.

There are no open issues under Measurement 44.

I. Other Issues.

The Settling Parties propose the following *additional* modifications to OSS performance measurements and standards that affect multiple measurements:

a. For maintenance measures for DSL (including Line Sharing), Verizon will provide separate disaggregation for UNE loops meeting standard criteria for DSL services and UNE loops that do not meet standard criteria. They propose that performance be assessed for standard UNE loops and tracked diagnostically for non-standard UNE loops.

b. They propose certain clarifications to Verizon's definitions of service group types and respective analogs.

APPENDIX B
Page 39

c. They propose to measure Pacific's Optical Carrier (OC) level services, including Enhanced Extended Links (EELs) as separate service group types.

d. They propose that Pacific's report date be moved from the 15th of the month to the 20th day of the month.

e. They propose adding language under the “Reporting Process” section which describe Pacific’s commitments to reporting on the 20th day of the month, instead of the 15th.

f. They also propose replacing Verizon’s jeopardy codes with new codes.

The Settling Parties continue to disagree about the following issues:

a. A proposal to evaluate performance results for Pacific's / Verizon's data affiliates against the better of parity or benchmark.

b. A proposal to establish an interim benchmark for all measures that show xDSL as a parity measurement of Verizon's separate data affiliate (SDA), which is not yet operational.

c. A proposal to move Verizon's reporting date from the 15th of the month to the 20th of the month.

d. A proposal that Pacific provide separate disaggregation for UNE loops meeting standard criteria for DSL services and UNE loops that do not meet standard criteria. Nor do they agree that Pacific's performance will be assessed for standard UNE loops and tracked diagnostically for non-standard UNE loops.

The Settling Parties have submitted the aforementioned disputes for resolution by the Commission.

(END OF APPENDIX B)

APPENDIX C

Revised
As Of 10/27/00

California OSS OII Performance Measurements

Joint Partial Settlement Agreement

INTRODUCTION

On October 9, 1997, the Commission issued an order instituting a rulemaking proceeding and investigation (hereinafter, the "OSS OII") to accomplish several goals, including the determination of reasonable standards of OSS performance for Pacific and GTE, the development of a mechanism that will allow the Commission to monitor improvements in OSS performance, and the assessment of the best and fastest method of ensuring compliance if standards are not met, or improvement is not shown[1].

Pursuant to the Commission's issuance of the OSS OII, the Settling Parties entered into lengthy and detailed negotiations to establish a set of performance measures consistent with the Commission's stated goals.[1] The Settling Parties filed a Joint Motion for approval of the JPSA on January 7, 1999, and filed motions on the remaining open issues on January 8, 1999. The Commission issued a decision approving the JPSA and resolving most of the remaining open issues on August 5, 1999. D.99-08-020.

The JPSA, as approved by the Commission in August 1999, called for a periodic review commencing in February 2000. Numerous meetings were held between the ILECs and CLECs to negotiate and resolve issues that have arisen over the past year. This iteration of the JPSA is a direct result of those collaborative sessions.

The issue of performance incentives is pending before the Commission.

The Commission staff has strongly encouraged CLECs and ILECs to stipulate to a resolution in this proceeding. This partial settlement agreement represents such a stipulation by the parties. This partial settlement report addresses the following:

- the performance measurements
- the formulas for the same
- the levels of disaggregation
- the analogs for the service group types (a level of disaggregation)
- other analogs and the benchmarks
- auditing and reporting
- review procedures

[1] A full history of the parties' negotiations and the basis for the development of the measures and standards contained in the JPSA is set forth in the Settling Parties' Joint Motion filed in this docket on January 7, 1999, and is incorporated by reference herein.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Performance Measures Development Process

The Telecommunications Act of 1996 and the FCC's implementing rules require Pacific and GTEC to provide CLECs with nondiscriminatory access to OSS. In the August 1996 Local Competition First Report and Order, the FCC commented, generally, that ILECs must provide CLECs with access to the pre-ordering, ordering, provisioning, billing, repair, and maintenance OSS sub-functions pursuant to the Act such that CLECs are able to perform such OSS sub-functions in "substantially the same time and manner" as the ILECs can for themselves[2]. The FCC's 271 decisions have analyzed the nondiscriminatory access requirements of§251(c) to a Bell Operating Company's (BOC's) §271 application, and clarified that for those OSS subfunctions with retail analogs, a BOC "must provide access to competing carriers that is equal to the level of access that the BOC provides to itself, its customers or its affiliates, in terms of quality, accuracy and timeliness."[3] The FCC further clarified that for those OSS functions with no retail analog, a BOC must offer access sufficient to allow an efficient competitor "a meaningful opportunity to compete."[4]

Initially, some of the interconnection agreements contained performance measures. In late 1997, the California Public Utilities Commission (CPUC) initiated OSS OII/OIR Docket 97-10-016 and 97-10-017 to address monitoring the performance of Operations Support Systems (OSS). The three stated goals of the Commission's OSS/OII proceeding are:

---

[2] See, Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, CC Docket No. 96-98, First Report and Order, 11 FCC Rcd 15499, 15763-64 [¶518] (1996) ("Local Competition First Report and Order"), aff'd in part and vacated in part sub nom. Competitive Telecommunications Ass'n v. FCC, 117 F.3d 1068 (8th Cir. 1997) and Iowa Utilities Bd. v. FCC, 120 F.3d 753 (8th Cir. 1997), modified on reh'g, No. 96-3321 (Oct. 14, 1997) (Rehearing Order), petition for cert. granted, 118 S. Ct. 879 (1998).

[3] *See In the Matter of Application by Bell Atlantic New York for Authorization Under Section 271 of the Communications Act to Provide In-Region , InterLATA Service in the State of New York, CC Docket No.99-295. See also, In the Matter of Application of Ameritech Michigan Pursuant to Section 271 of the Communications Act of 1934, as amended, To Provide In-Region, InterLATA Services In Michigan,* Memorandum Opinion and Order, 12 FCC Rcd 20543, 20618-19 [¶139] (1997) (*Ameritech Michigan Order*), *writ of mandamus issued sub nom. Iowa Utils. Bd. v. FCC,* No. 96-3321 (8th Cir. Jan. 22, 1998). ("*Ameritech Opinion*"); *see also, In the Matter of Application of Bellsouth Corporation, et al., for Provision of In-Region, InterLATA services in Louisiana ("BellSouth (Louisiana II) Opinion")* CC Docket No. 98-121, FCC 98-271 (10-13-98), paragraph 87 (*citing, Ameritech Opinion* at 12 FCC Rcd 20618-19). *See also, Ameritech Opinion* at ¶131, wherein the FCC makes the following statement regarding application of the §251(c) requirements to a BOC's §271 application:

> "Because the duty to provide access to network elements under section 251(c)(3) and the duty to provide resale services under section 251(c)(4) include the duty to provide nondiscriminatory access to OSS functions, an examination of a BOC's OSS performance is necessary to evaluate compliance with section 271(c)(2)(B)(ii) and (xiv)."

[4] *See In the Matter of Application by Bell Atlantic New York for Authorization Under Section 271 of the Communications Act to Provide In-Region , InterLATA Service in the State of New York, CC Docket No.99-295. See also, Ameritech Opinion* at 12 FCC Rcd at 20619 [¶141]; *See also, BellSouth (Louisiana II) Opinion* at ¶87 (*citing Ameritech Opinion* at 12 FCC Rcd at 20619).

- "to determine reasonable standards of performance for Pacific Bell (Pacific) and GTE California Incorporated (GTEC) in their Operations Support Systems (OSS),
- to develop a mechanism that will allow the Commission to monitor improvements in the performance of OSS, and
- to assess the best and fastest method of ensuring compliance if standards are not met or improvement is not shown. A subset of the third goal will be to provide appropriate compliance incentives under Section 271 of the Telecommunications Act of 1996, which applies solely to Pacific for the prompt achievement of OSS improvements."[5]

The scope of the proceeding included measures, reporting, comparative analogs, benchmarks, statistical tests, audits and incentives. This report is not intended to address statistical tests and incentives.

Major Categories

Measurements developed to help assess the provision of non-discriminatory access to OSS and other services, elements or functions were combined into the following broad categories:

- **Pre-Ordering**

Pre-ordering activities relate to the exchange of information between the ILEC and the CLEC regarding current or proposed customer products and services, or any other information required to initiate ordering of service. Pre-ordering encompasses the critical information needed to submit a provisioning order from the CLEC to the ILEC. The pre-order measurement reports the timeliness with which pre-order inquiries are returned to CLECs by the ILEC. Pre-ordering query types include:

Address Verification/Dispatch Required
Request for Telephone Number
Request for Customer Service Record
Service Availability
Service Appointment Scheduling (due date)
Loop Qualification
Facility Availability
Rejected/Failed Inquiries

- **Ordering**

Ordering activities include the exchange of information between the ILEC and the CLEC regarding requests for service. Ordering includes: (1) the submittal of the service request from the CLEC, (2) rejection of any service request with errors and (3) confirmation that a valid service request has

[5] Order Instituting Rulemaking on the Commission's Own Motion into Monitoring Performance of Operations Support Systems (R.97-10-016), and Order Instituting Investigation on the Commission's Own Motion into Monitoring Performance of Operations Support Systems (I.97-10-017), October 9, 1997.

been received and a due date for the request assigned. Ordering performance measurements report on the timeliness with which these various activities are completed by the ILEC. Also captured within this category is reporting on the number of CLEC service requests that automatically generate a service order in the ILECs' service order creation system.

- **Provisioning**

Provisioning is the set of activities required to install, change or disconnect a customer's service. It includes the functions to establish or condition physical facilities as well as the completion of any required software translations to define the feature functionality of the service. Provisioning also involves communication between the CLEC and the ILEC on the status of a service order, including any delay in meeting the commitment date and the time at which actual completion of service installation has occurred. Measurements in this category evaluate the quality of service installations, the efficiency of the installation process and the timeliness of notifications to the CLEC that installation is completed or has been delayed.

- **Maintenance**

Maintenance involves the repair and restoral of customer service. Maintenance functions include the exchange of information between the ILEC and CLEC related to service repair requests, the processing of trouble ticket requests by the ILEC, actual service restoral and tracking of maintenance history. Maintenance measures track the timeliness with which trouble requests are handled by the ILEC and the effectiveness and quality of the service restoral process.

- **Network Performance**

Network performance involves the level at which the ILEC provides services and facilitates call processing within its network. The ILEC also has the responsibility to complete network upgrades efficiently.. Network performance is evaluated on the quality of interconnection and the timeliness of network upgrades (code openings) the ILEC completes on behalf of the CLEC.

- **Billing**

Billing involves the exchange of information necessary for CLECs to bill their customers, to process the end user's claims and adjustments, to verify the ILEC's bill for services provided to the CLEC and to allow CLECs to bill for access. Billing measures have been designed to gauge the quality, timeliness and overall effectiveness of the ILEC billing processes associated with CLEC customers.

- **Collocation**

ILECs are required to provide to CLECs available space as required by law to allow the installation of CLEC equipment. Performance measures in this category assess the timeliness with which the ILEC handles the CLEC's request for collocation as well as how timely the collocation arrangement is provided.

- **Data Base Updates**

Database updates for directory assistance/listings and E911 include the processes by which these systems are updated with customer information which has changed due to the service provisioning activity. Measurements in this category are designed to evaluate the timeliness and accuracy with which changes to customer information, as submitted to these databases, are completed by the ILEC.

- **Interfaces**

ILECs provide the CLECs with choices for access to OSS pre-ordering, ordering, maintenance and repair systems. Availability of the interfaces is fundamental to the CLEC being able to effectively do business with the ILEC. Additionally, in many instances, CLEC personnel must work with the service personnel of the ILEC. Measurements in this category assess the availability to the CLECs of systems and personnel at the ILEC work centers.

Auditing and Review Procedures

The parties have agreed to the procedures for auditing and review. Descriptions of these procedures can be found in Sections IV and V.

Note: This Executive Summary is intended to provide a general background regarding parties' negotiations of the OSS performance measures. The statements contained in the Executive Summary are not intended to be legally binding on the parties and shall not be used for such purposes.

Reservation of Rights

These reservations of rights do not negate the parties agreement regarding performance measures and standards as reflected in this settlement agreement.

Incorporating the performance measures into the interconnection agreements raises several complex issues. The Commission has indicated it will rule on this matter in a subsequent decision.

ILECs

By agreeing to the performance measures contained in the Joint Partial Settlement Agreement, ILECs**:**

- do not make any admission regarding the propriety or reasonableness of establishing performance penalties;
- reserve the right to contest the level of disaggregation for purpose of assessing penalties;
- reserve the right to contend that any resulting penalties should viewed as liquidated damages and as the exclusive remedy for any failure of performance; and,
- do not admit that an apparent less-than-parity condition reflects discriminatory treatment without further factual analysis.

CLECs

- By executing this Agreement, CLECs do not agree with, endorse, or otherwise concur in the terms of ILECs' reservation of rights.
- CLECs reserve the right to contend that ILEC compliance with the performance measures and standards in the Agreement does not conclusively demonstrate ILEC compliance with the Telecommunications Act of 1996.
- CLECs reserve the right to contend that ILEC compliance with the performance measures and standards does not conclusively demonstrate the existence of an open competitive local market.

NOTES:

1. *Not all measures apply to both ILECs*.

2. *These performance measures are not intended to create, modify or otherwise affect parties' rights and obligations. The existence of any particular performance measure, or the language describing that measure, is not evidence that the CLECs are entitled to any particular manner of access, that these measures relate solely to access to OSS, or is it evidence that the ILEC's obligations are limited to providing any particular manner of access. The parties' rights and obligations to such access are defined elsewhere, including the relevant laws, FCC and CPUC decisions/regulations, tariffs, and interconnection agreements.*

3. *Details regarding implementation schedules for new measures are documented in Section VI (Implementation Schedules).*

Pre-Ordering **Measure 1**

Title: Average Response Time (to Pre-Order Queries)

Area	*Requirement Description*

Description:	This measure captures the response interval for each pre-ordering query. It is determined by computing the elapsed time from the ILEC receipt of the query from the CLEC, whether or not syntactically correct, to the time the ILEC returns the requested data to the CLEC.
	• Address Verification/Dispatch Required
	• Request for Telephone Number
	• Request for Customer Service Record
	• Service Availability
	• Service Appointment Scheduling (due date)
	• Rejected/Failed inquires
	• Facility Availability (Pacific Bell Only)
	• Loop qualification
	    • Loop Qual (Mechanized)
	    • K1023 loop qualification (Pacific Bell)
	        • xDSL and High Bandwidth line sharing UNE loop qualification
	        • All Other loop qualification

Method of Calculation:	**Mechanized:** **<u>Pre - Order Query Transaction Time</u>** Sum ((Query Response Date and Time) – (Query Submission Date and Time)) / (Number of Queries Returned in Reporting Period) **<u>Legacy System Transaction Time (GTE only)</u>** Sum ((Query Response Date and Time from Legacy System) – (Query Submission Date and Time to Legacy System)) / (Number of Queries Returned to Legacy System in Reporting Period) **<u>Loop Qualification/Facility Availability Transaction Time (Pacific Bell Only)</u>** Sum ((Query Response Date and Time) – (Query Submission Date and Time)) / (Number of Queries Returned in Reporting Period) **<u>Loop Qualification Transaction Time (GTE Only)</u>** Sum ((Query Response Date and Time) - (Query Submission Date and Time)) / (Number of Queries Returned in Reporting Period) **<u>Manual CSRs (Pacific Bell and GTE)</u>** (# of CSR's Returned within "X" Business Hours) / (# of CSRs Returned) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and ILEC affiliate
Reported By:	By query type and by interface type, including fax
Geographic Level:	Statewide

Measurable Standard:	**Mechanized:** Standard: | Pacific Bell | GTE Address Verification | av. 4.5 seconds | Legacy Time + 5 seconds TN Selection | av. 4.5 seconds | Legacy Time + 5 seconds CSR | av.10.0 seconds | 98% within 3 hrs. (WISE) TBD (EDI/CORBA) Service Availability | av. 8.0 seconds | Legacy Time + 5 seconds Due Date | av. 2.0 seconds | Legacy Time + 5 seconds Reject/Failed Inquiries Dispatch | av. 11.0 seconds | N/A (Inc. in Address Verification) **Manual CSRs:** **Pacific Bell:** **Benchmark:** • Standard - 95% in 4 hours **GTE:** **Benchmark:** • Standard - 98% in 24 hours **Mechanized Loop Qualification:** • Standard - Parity (Pacific Bell) • Standard - Benchmark - TBD (GTE) **Manual Loop Qualification (K1023) Process (Pacific Bell only)** • Standard - Parity

Business Rules:	• Pre-order query transaction time intervals are measured as total transaction time. • For Pacific Bell, excludes CSR requests (both manual and mechanized) for greater than 50 working telephone numbers • For Pacific Bell, fully electronic pre-order query response times will be measured for the Verigate, Datagate and Loop Qual systems. Pre-ordering functionality only recently made available for EDI/CORBA. Benchmarks will be established by November 15, 2000. • For GTE fully electronic pre-order query response times will be measured for the WISE and CORBA systems. • For GTE, manual CSRs measured in clock hours; excludes non-business days. • Elapsed time for fully electronic sub-measures tracked during published system hours. • Mechanized Loop Qualification measured in seconds. (Pacific Bell only) • Elapsed time for manual processes tracked during published business hours.(Pacific Bell only) • Response time for Pacific Bell's Starwriter system is measured at parity based on % within 4 seconds. • GTE does not report Legacy System Transaction Time for rejected/failed inquiries. • Pre-Order Query Transaction Time will be reported and tracked diagnostically for rejected/failed inquiries.
Notes:	• The numerator and denominator of the sub-measures in this measure capture all queries completed in the reporting period. • GTE will supply all available loop qualification data, however GTE will not support manual engineering query for loop qualification. • Where CLEC accesses Pacific Bell's systems using a Service Bureau Provider, the measurement of Pacific Bell's performance shall not include the Service Bureau Provider's processing, availability or response time. •

OSS OII Performance Measurements
Report Requirements

Ordering — Measure 2

Title: Average FOC/LSC Notice Interval

Area	*Requirement Description*
Description:	Measures the average time from receipt of a valid service request to returning a Firm Order Confirmation (FOC)/Local Service Confirmation (LSC).
Method of Calculation:	**Mechanized:** Sum ((Date and Time of FOC/LSC) - (Business Date and Time of Receipt of Valid Service Request)) / (Number of FOCs/LSCs Sent in Reporting Period) **Manual:** Sum ((Fax Date and Time Returned) - (Business Date and Time receipt of valid fax service request)) / (Number of Faxes Submitted in Reporting period) **Held and Denied Interconnection Trunk Requests:** [(Sum (Date Request is Released) – (Date Request is Originally Received)]/ (Number of Requests Held and Released)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and ILEC affiliates.
Reported By:	• Electronically received/electronically handled • Electronically received and manually handled • Manually received and manually handled • By service group type and Stand Alone Directory Listings (GTE only)
Geographic Level:	Statewide

Measurable Standard:	**Service Group Types:**
	Pacific Bell
	• Resale Residential POTS
	• Resale Business POTS
	• Resale ISDN BRI
	• Resale CENTREX
	• Resale PBX
	• Resale DDS
	• Resale DS1/ISDN-PRI
	• Resale DS3
	• Resale VGPL/DS0
	• 2/4w (8db) analog loop (incl. Coin/analog PBX)
	• 2w digital loop(ISDN capable)
	• 2w digital loop(xDSL capable)
	• High Bandwidth Line Sharing UNE
	• 4w digital loop DS1
	• UNE loop – DS3
	• UNE Loop – OC level
	• UNE Dark Fiber
	• UNE Port– Non-Specials)
	• UNE Port–Specials
	• UNE Dedicated Transport
	    • DS1
	    • DS3
	    • OC level
	• Enhanced Extended Links
	    • VG
	    • DS1
	    • DS3
	    • OC level
	• UNE Platform
	    • Basic port and loop
	    • Special port and basic loop
	    • ISDN BRI port and loop
	    • ISDN PRI port and loop
	• Standalone LNP
	• Interconnection Trunks
		GTE
		• Resale POTS- Residence
		• Resale POTS-Business
		• Resale Specials
		• UNE loop Nondesigned
		• UNE loop Designed
		• UNE loop xDSL capable
		• UNE loop IDSL capable
		• UNE Port
		• UNE Transport
		• UNE Platform
		• UNE-P Res
		• UNE-P Bus
		• UNE-P PRI
		• Interconnection Trunks
		• Line Sharing - Conditioned
		• Line Sharing - Non -Conditioned
		• LNP
		• EEL (Diagnostic)
		• Subloop (Diagnostic)
		• Dark Fiber (Diagnostic)

Measurable Standard:	**Benchmark:** **Fully Electronic/Flow Through:** • Standard - average of 20 minutes **Electronically Received/Manually Handled** • Standard - average of 6 hours **Manually received/Manually Handled** • Standard - average of 12 hours **Projects:** • Standard -90% within 72 hours (Pacific Bell) **Interconnection Trunks** • Standard: **Pacific Bell:** **GTE:** Average 7 business days (New) Average 5 business day (All) Average 4 business days (Augment) **Interconnection Trunk Requests:** **Held and Denied –** Average Interval • Standard - Parity (Pacific Bell only) • Standard – Average 13 days (GTE only) .

Business Rules:	• The start time of requests received after the end of the business day will be the beginning of the next business day. Business day is defined as published hours of operation for the ILEC ordering center.
	    • Business day = Monday through Friday, excluding weekends and ILEC published holidays
	• Excludes non-business days.
	• Excludes delays caused for customer reasons
	• Elapsed time for fully electronic sub-measures tracked during system hours.
	• Loop qualification/availability of facilities interval is excluded from overall FOC interval for the following products: (Pacific Bell only)
	    • xDSL and High Bandwidth line sharing UNE
	    • ISDN
	    • Channelized DS1
	    • DS3
	    • Dark Fiber
	    • Unbundled Dedicated Transport - DS3
	• ILEC will only perform pre-qualification for above mentioned UNEs if pre-qualification has not been completed prior to the submission of the service request by the CLEC, and it is required
	• Projects are defined as POTS greater than 20 lines, for Specials greater than 6 lines, UNE Loops greater than 20 loops, and Interconnection Trunks greater than 192 trunks.(Pacific Bell only)

Notes:	• Where CLEC accesses Pacific Bell's systems using a Service Bureau Provider, the measurement of Pacific Bell's performance shall not include the Service Bureau Provider's processing, availability or response time.

OSS OII Performance Measurements
Report Requirements

Ordering Measure 3

Title: Average Reject Notice Interval

Area	*Requirement Description*
Description:	Reject interval is the elapsed time between the ILEC receipt of an order from the CLEC to the ILEC return of a notice of a rejection to the CLEC.
Method of Calculation:	**Mechanized:** Sum ((Business Date and Time of ILEC Transmission of Order Rejection) - (Business Date and Time of Order Receipt)) / (Number of MechanizedOrders Rejected in the Reporting Period) **Manual:** Sum ((Fax Date and Time Returned) - (Business Date and Time Receipt of fax service request)) / (Number of Faxes Rejected in Reporting Period)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and ILEC Affiliates

Reported By:	• Electronically received, electronically handled
	  • All interfaces
	  • Syntax(edit engine) and content errors (other edits)
	  • Resale orders, High Bandwidth line sharing UNE, other Facility based/UNE orders and standalone Directory Listings
	• Electronically received, manually handled
	  • All interfaces
	  • Syntax (edit engine) and content errors (other edits)
	  • Resale orders, High Bandwidth line sharing UNE and other Facility based/UNE orders and standalone Directory Listings (GTE only)
	• Manually received and handled (fax)
	  • Resale orders, High Bandwidth line sharing UNE and other Facility based/UNE orders and standalone Directory Listings (GTE only)

Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell and GTE:** **Benchmark:** **Fully Electronic/Flow Through:** • Standard - average of 20 minutes **Electronically Received/Manually Handled:** • Standard - average of 5 hours **Manually received/Manually Handled:** • Standard - average of 10 hours **Projects:** • Standard -90% within 72 hours (Pacific Bell only)

Business Rules:	• Elapsed time for fully electronic sub-measures tracked during system hours
	• For manually handled requests:
	Calculation of requests received after the end of the business day starts at the beginning of the next business day. Business day is defined as published hours of operation for the ILEC.
	• Business day = Monday through Friday, excluding weekends and ILEC published holidays
	• • Excludes non-business days
	• Excludes delays caused for customer reasons
	• Loop qualification/facility availability interval is removed from the overall reject interval for the following products: (Pacific Bell only)
	  • XDSL
	  • High Bandwidth line sharing UNE
	  • ISDN
	  • Channelized DS1
	  • DS3
	  • Dark Fiber
	  • Unbundled Dedicated Transport - DS 3
	• ILEC will only perform pre-qualification for above mentioned UNEs if pre-qualification has not been completed prior to the submission of the service request by the CLEC, and it is required.
	• Projects are defined as POTS greater than 20 lines, for Specials greater than 6 lines, UNE Loops greater than 20 loops, and Interconnection Trunks greater than 192 trunks.(Pacific Bell only)

Notes:	• All benchmarks adopted are interim: the parties should collect data and submit proposed modifications of the adopted measurable standards by February 1, 2000(Benchmarks for GTE are still interim.) • Where CLEC accesses Pacific Bell's systems using a Service Bureau Provider, the measurement of Pacific Bell's performance shall not include the Service Bureau Provider's processing, availability or response time.

Ordering **Measure 4**

Title: Percentage of Flow-Through Orders

Area	*Requirement Description*
Description:	Measures the percentage of electronically received orders processed on a flow through basis.
Method of Calculation:	[(Number of valid electronically received orders that flow-through without manual intervention) / (Total valid electronically received orders)] x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, and ILEC Affiliates
Reported By:	Orders that flow through as a percentage of: • All electronically received orders programmed to flow through, by service group type and/or service order type. • All electronically received orders, by service group type and/or service order type.
Geographic Level:	Statewide
Measurable Standard:	**Diagnostic only** *Issue of how to evaluate performance will be reconsidered at next Performance Measurement Plan review.*
Business Rules:	• Excludes orders rejected due to CLEC caused syntax errors, but does not exclude CLEC caused content errors.
Notes:	

Provisioning **Measure 5**

Title: Percentage of Orders Jeopardized

Area	*Requirement Description*
Description:	Percentage of total orders processed for which the ILEC notifies the CLEC that the work will not be completed as committed on the original FOC.
Method of Calculation:	((Number of Orders Jeopardized) / (Number of Orders Confirmed)) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies)and ILEC Affiliates
Reported By:	• By service group type
Geographic Level:	Statewide

Measurable Standard:	**<u>Pacific Bell:</u>** **Parity for Resale is Retail Parity measured** **for the following UNEs:**	**Retail**
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX) • UNE Subloop	• POTS - Business (fielded)
	• 2w digital loop(ISDN capable) • UNE Subloop	• ISDN(BRI)
	• 2w digital loop(xDSL capable) • UNE Subloop	• 2w digital loop(xDSL capable) provided to ASI
	• 2w digital loop(IDSL capable) • UNE Subloop	• ISDN(BRI)

	• High Bandwidth Line Sharing UNE
	    • Conditioned
	    • Non-Conditioned
		• High Bandwidth Line Sharing UNE provided to ASI

	• 4w digital loop (DS1) • UNE Subloop	• DS1
	• UNE loop – DS3	• DS3
	• UNE Loop – OC level	• Retail OC level service
	• Dark Fiber	(*Diagnostic*)
	• UNE Port–(Non-Specials)	• POTS - Business (non-fielded)
	• UNE Port–Specials	• Retail Specials (non-fielded)

	• UNE Dedicated Transport
	    • DS1
	    • DS3
	    • OC level
		• HICAP
		    • DS1
		    • DS3
		    • Retail OC level service

	• Enhanced Extended Links
	    • VG - Conversion
	    • DS1 - New
	    • DS1 -Conversion
	    • DS3- New
	    • DS3-Conversion
	    • OC level – New
	    • OC level - Conversion
		(*TBD*)

	• UNE Platform
	    • Basic port and loop
	    • Special port and basic loop
	    • ISDN BRI port and loop
	    • ISDN PRI port and loop
		• Business POTS FW/NFW
		• Retail Voice Grade Specials FW/NFW
		• ISDN BRI FW/NFW
		• ISDN PRI FW/NFW

	• Interconnection Trunks	• ILEC Dedicated Trunks

Measurable Standard:	**GTE**	**Retail**
	• Resale POTS- Residence	• Retail POTS - Residence
	• Resale POTS-Business	• Retail POTS - Business
	• Resale Specials	• Retail Specials
	• UNE loop Nondesigned	• B1 Dispatched Non Designed
	• UNE loop Designed	• Dispatched Designed Service (excludes HICAPs)
	• UNE loop xDSL capable	• *(TBD until SDA is established)*
	• UNE Loop IDSL capable	• *(TBD until SDA is established)*
	• UNE Port	• CentraNet - Simple
	• UNE Transport	• HICAP Designed
	• UNE Platform	
	• UNE-P Res	• Retail POTS
	• UNE-P Bus	• Business POTS
	• UNE-P PRI	• ISDN PRI
	• Interconnection Trunks	• ILEC Dedicated Trunks
	• Line Sharing - Conditioned	• *(TBD until SDA is established)*
	• Line Sharing - Non Conditioned	• *(TBD until SDA is established)*
	• LNP	• Retail POTS -*Total Business & Residence, Non-Dispatched*
	• EEL	• *(Diagnostic)*
	• Subloop	• *(Diagnostic)*
	• Dark Fiber	• *(Diagnostic)*
Business Rules:	• Excludes delays for customer reasons. • Raw data will include jeopardy codes. • For Pacific Bell results for UNE Subloop will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity with ASI • For GTE results for UNE subloop will be tracked diagnostically. • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review	
Notes:	• Does not include missed commitments.	

Provisioning **Measure 6**

Title: Average Jeopardy Notice Interval

Area	*Requirement Description*
Description:	Measures the remaining time between the pre-existing committed order completion date and time (communicated via the FOC) and the date and time the ILEC issues a notice to the CLEC indicating an order is in jeopardy of missing the due date (or the due date/time has been missed).
Method of Calculation:	**Assignment:** *Jeopardies identified during the initial assignment process* Sum ((Date of Committed Due Date for the Order) - (Date of Jeopardy Notice)) / (Number of Assignment Jeopardy Notices) **Installation:** *Jeopardies identified during the installation process prior to due time* Sum ((Date & Time of Committed Due Date for the Order) - (Date & Time of Jeopardy Notice)) / (Number of Installation Jeopardy Notices) **Notification of Missed Commitments** Sum(Due Date and Time of Missed Commit Notice - Due Date and Time of Order) / (Number of Missed Commit Notices)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, and ILEC Affiliates
Reported By:	• By service group type, with same service group type disaggregation as Measure 5.
Geographic Level:	Statewide

Measurable Standard:	**Service Group Types:**
	Pacific Bell
	• Resale Residential POTS
	• Resale Business POTS
	• Resale ISDN BRI
	• Resale CENTREX
	• Resale PBX
	• Resale DDS
	• Resale DS1/ISDN-PRI
	• Resale DS3
	• Resale VGPL/DS0
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX)
	    • UNE Subloop
	• 2w digital loop(ISDN capable)
	    • UNE Subloop
	• 2w digital loop(xDSL capable)
	    • UNE Subloop
	• High Bandwidth Line Sharing UNE
	    • Conditioned
	    • Non-Conditioned
	• 4w digital loop DS1
	    • UNE Subloop
	• UNE Loop – DS3
	• UNE Loop –OC level
	• UNE Dark Fiber
	• UNE Port– Non-Specials
	• UNE Port–Specials
	• UNE Dedicated Transport
	    • DS1
	    • DS3
	    • OC level
	• Enhanced Extended Links
	    • VG - Conversion
	    • DS1 - New
	    • DS1 - Conversion
	    • DS3 -New
	    • DS3 - Conversion
	    • OC Level – new
	    • OC level - conversion
	• UNE Platform
	    • Basic port and loop
	    • Special port and basic loop
	    • ISDN BRI port and loop
	    • ISDN PRI port and loop
	• Interconnection Trunks
		GTE
		• Resale POTS- Residence
		• Resale POTS-Business
		• Resale Specials
		• UNE loop Nondesigned
		• UNE loop Designed
		• UNE loop xDSL capable
		• UNE loop IDSL capable
		• UNE Port
		• UNE Transport
		• UNE Platform
		    • UNE-P Res
		    • UNE-P Bus
		    • UNE-P PRI
		• Interconnection Trunks
		• Line Sharing - Conditioned
		• Line Sharing - Non -Conditioned
		• LNP
		• EEL (Diagnostic)
		• Subloop (Diagnostic)
		• Dark Fiber (Diagnostic)

Measurable Standard:	**Benchmark (Pacific Bell only)** • Standard - Assignment Jeopardies 90% within 1 day Install. Jeopardies (POTS) 95% within 15 minutes Install. Jeopardies (Specials) 95% within 3 hours Missed Commit Notices 95% within 24 hours *GTE began reporting June 2000 data on July 15, 2000. GTE will propose benchmark after four months of data collection.*
Business Rules:	• Excludes delays for customer reasons. • Raw data will include jeopardy codes. • Pacific Bell tracks assignment jeopardies by due date only, installation jeopardies by business days/hours and notifications of missed commitments by clock hours. • GTE tracks assignment jeopardies by due date only for business days, with installation jeopardies and notifications of missed commitments tracked by business days/hours.
Notes:	• If the ILECs' policy regarding jeopardy notices to their Retail customers changes, this measure should be evaluated for analog. • For GTE, jeopardies issued on the due date are considered either installation or notifications of missed commitments.

Provisioning — Measure 7

Title: Average Completed Interval

Area	*Requirement Description*
Description:	Average business days from receipt of valid, error-free service request to completion date in service order system for new, move, and change orders.
Method of Calculation:	Total business days from receipt of valid, error-free service request to completion date in service order system for new, move and change orders / Total new, move and change orders
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and ILEC Affiliates
Reported By:	By service group type and field work/no field work where applicable.
Geographic Level:	Region (PB), Statewide (GTE)

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail for** **Parity for UNE measured** **for the following UNEs:**	**Retail**
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX) ◦ UNE Subloop	• POTS - Business (fielded)
	• 2w digital loop(ISDN capable) ◦ UNE Subloop	• ISDN(BRI)

	• 2w digital loop(xDSL capable)
	  ◦ Conditioned
	  ◦ Non-Conditioned
	  ◦ UNE Subloop
		• 2w digital loop (xDSL capable) provided to ASI
		  ◦ Conditioned
		  ◦ Non-Conditioned

	• 2w digital loop(IDSL capable) ◦ UNE Subloop	• ISDN(BRI)

	• High Bandwidth line sharing
	  ◦ Conditioned
	  ◦ Non-Conditioned
		• High Bandwidth line sharing provided to ASI
		  ◦ Conditioned
		  ◦ Non-Conditioned

	• 4w digital loop (DS1)	• DS1
	• UNE Loop – OC level	• Retail – OC level service
	• UNE Port– Non-Specials	• POTS - Business (non -fielded)
	• UNE Port–Specials	• Retail Special Services

	• UNE Dedicated Transport
	  ◦ DS1
	  ◦ DS3
	  ◦ OC level
		• HICAP
		  ◦ DS1
		  ◦ DS3
		  ◦ Retail OC level service

	• Dark Fiber	(Diagnostic)

	• Enhanced Extended Links
	  ◦ VG - Conversion
	  ◦ DS1 - New
	  ◦ DS1 -Conversion
	  ◦ DS3- New
	  ◦ DS3-Conversion
	  ◦ OC level – New
	  ◦ OC level - Conversion
		(TBD)

	• UNE Platform • Basic port and loop • Special port and basic loop • ISDN BRI port and loop • ISDN PRI port and loop	• Business POTS FW/NFW • Retail Voice Grade Specials FW/NFW • ISDN BRI FW/NFW • ISDN PRI FW/NFW
	• Interconnection Trunks	• ILEC Dedicated Trunks

<table>
<tr><td rowspan="2">Measurable Standard:</td><td><u>GTE</u></td><td>Retail</td></tr>
<tr><td>
• Resale POTS- Residence
• Resale POTS-Business
• Resale Specials
• UNE loop Nondesigned
• UNE loop Designed

• UNE loop xDSL capable
• UNE loop IDSL capable
• UNE Port
• UNE Transport
• UNE Platform
  • UNE-P Res
  • UNE-P Bus
  • UNE-P PRI
• Interconnection Trunks
• Line Sharing - Conditioned
• Line Sharing - Non -Conditioned
• LNP

• EEL
• Subloop
• Dark Fiber
</td><td>
• Retail POTS - Residence
• Retail POTS - Business
• Retail Specials
• B1 Dispatched Non Designed
• Dispatched Designed Service (excludes HICAPs)
• (TBD until SDA is established)
• (TBD until SDA is established)
• CentraNet-Simple
• HICAP Designed

• Residential POTS
• Business POTS
• ISDN PRI
• ILEC Dedicated Trunks
• (TBD until SDA is established)
• (TBD until SDA is established)
• Retail POTS -Total Business & Residence, Non-Dispatched
• (Diagnostic)
• (Diagnostic)
• (Diagnostic)
</td></tr>
</table>

Business Rules:	• Excludes customer requested due dates other than interval offered, and orders delayed for customer reasons. (Pacific Bell only) • Excludes customer due dates beyond interval offered, and orders delayed for customer reasons. (GTE) • For UNE loop services, feature-only orders are excluded from retail analog.(Pacific Bell only) • Excludes projects. (Pacific Bell only) • GTE will not exclude projects. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity with ASI (Pacific Bell only) • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review. • The Completion Date is the date on which the service has passed acceptance testing, where applicable. To the extent that Pacific is required to obtain affirmative acceptance of the loop from the CLEC before closing an order, the order will not be deemed to have successfully passed an acceptance test until the CLEC affirmatively accepts the loop. (Pacific Bell only) • Orders where acceptance testing is delayed as a result of CLEC action or inaction shall be excluded. (Pacific Bell only)
Notes:	• For Pacific Bell, no retail analog exists for IDSL capable loops. The retail comparison will be made with ISDN service which has similar characteristics.

Provisioning

Measure 8

Title: Percent Completed Within Standard Interval

Area	*Requirement Description*
Description:	Measures of orders completed within the standard interval of receipt of valid, error-free service request.
Method of Calculation:	Sum (Total New, Move and Change Orders Completed Within the Standard interval of Receipt of Valid, Error-free Service Request) / (Total New, Move and Change Orders)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and ILEC Affiliates
Reported By:	By service group type excluding services with flexible due dates.
Geographic Level:	Region (PB), Statewide (GTE)

Measurable Standard:	**<u>Pacific Bell</u>** **Parity for Resale is Retail** **Parity for UNE measured for the following UNEs:**	**Pacific Bell Retail**
	• 2w digital loop(ISDN capable) • UNE subloop	• ISDN(BRI)

	• 2w digital loop(xDSL capable)
	  • Conditioned
	  • Non-Conditioned
	  • UNE subloop
		• 2w digital loop (xDSL capable) provided to ASI
		  • Conditioned
		  • Non-Conditioned

	• 2w digital loop(IDSL capable) • UNE subloop	• ISDN (BRI)

	• High Bandwidth line sharing
	  • Conditioned
	  • Non-Conditioned
		• High Bandwidth line sharing provided to ASI
		  • Conditioned
		  • Non-Conditioned

	• 4w digital loop (DS1)	• DS1
	• UNE loop – OC level	• Retail – OC level service
	• Dark Fiber	*Diagnostic*
	• UNE Port– Specials	• Retail Specials

	• Enhanced Extended Links
	  • VG - Conversion
	  • DS1 - New
	  • DS1 -Conversion
	  • DS3- New
	  • DS3-Conversion
	  • OC level - New
	  • OC level -Conversion
		(*TBD)*

	• UNE Dedicated Transport
	  • . DS1
	  • DS3
	  • OC level
		• HICAP
		  • DS1
		  • DS3
		  • Retail OC level service

	UNE Platform • Special port and basic loop • ISDN BRI port and loop • ISDN PRI port and loop	• Retail Voice Grade Specials FW/NFW • ISDN BRI FW/NFW • ISDN PRI FW/NFW
	• Interconnection Trunks	• ILEC Dedicated Trunks
	GTE Resale Specials	Retail Specials

Business Rules:	• Excludes customer requested due dates other than the standard interval, and orders delayed for customer reasons. (Pacific Bell only) • Excludes customer requested due dates greater than the standard interval, and orders delayed for customer reasons. (GTE only) • Excludes services with flexible due date i.e., Basic Exchange services/POTS (Pacific Bell only) • For UNE loop services, feature-only orders are excluded from retail analog. (Pacific Bell only) • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity ASI. (Pacific Bell only). • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review. (Pacific Bell only) • The Completion Date is the date on which the service has passed acceptance testing, where applicable. To the extent that Pacific is required to obtain affirmative acceptance of the loop from the CLEC before closing an order, the order will not be deemed to have successfully passed an acceptance test until the CLEC affirmatively accepts the loop. (Pacific Bell only) • Orders where acceptance testing is delayed as a result of CLEC action or inaction shall be excluded. (Pacific Bell only)
Notes:	• For Pacific Bell, no retail analog exists for IDSL capable loops. The retail comparison will be made with ISDN service which has similar characteristics.

OSS OII Performance Measurements
Report Requirements

Provisioning **Measure 9**

Title: Coordinated Customer Conversion as a Percentage On-Time

Area	*Requirement Description*
Description:	**Pacific Bell:** Measures the percentage of coordinated cutovers (TBCC/CHC) completed by Committed time* where CLEC has requested coordination (including LNP). ** Note: "Committed time" means within one hour of committed order due time* **GTE:** Measures the percentage of coordinated orders completed by committed time* for all orders where CLEC has requested coordination (including LNP) **Note: "Committed time" means the actual conversion completion time is no greater than the committed completion interval plus one hour.*
Method of Calculation:	**Pacific Bell** ((Number of coordinated cutovers completed by committed time) / (Count of coordinated cutovers scheduled in reporting period)) x 100 **GTE** (Number of coordinated orders completed by committed due date and time) / (Count of coordinated orders completed in reporting period) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), by ILEC Affiliates
Reported By:	• Residence and Business conversions and LNP (PB only) • Coordinated Conversions and Coordinated Hot Cuts (GTE only)
Geographic Level:	Statewide

Measurable Standard:	**Parity for Pacific Bell:** **Pacific Bell Retail** **Coor. Conversions (Res.)** — Coor. Conv. -Res **Coor. Conversions (Bus.)** — Coor. Conv. -Bus **Coor. Conversions (LNP-Port Out**) — Coor. Conv. - - (LNP-Port In/Back) **Benchmark for GTE: 90% On Time** **Coordinated Conversion (CC)** ***Designed and Non-designed*** **<u>Line Size</u>** — **<u>Committed Completion Interval</u>** From 1 to 49 lines: — 1 Work Hour 50 to 99 lines: — 2 Work Hours 100 to 199 lines: — 3 Work Hours 200 plus lines: — 4 Work Hours **Coordinated Hot Cut (CHC)** **Designed and Non-designed** ***<u>Line Size</u>*** — ***<u>Committed Completion Interval</u>*** From 1 to 20 lines: — 1 Work Hour 21 to 30 lines: — 1½ Work Hours 31 to 40 lines: — 2 Work Hours 41 to 50 lines: — 2½ Work Hours 51 to 60 lines: — 3 Work Hours 61 to 70 lines: — 3½ Work Hours 71 to 80 lines: — 4 Work Hours 81 to 90 lines: — 4½ Work Hours 91 to 100 lines: — 5 Work Hours Add an additional ½ Hour for each additional 10 lines or increment thereof.
Business Rules:	• Excludes CLEC caused misses • Applies to CLEC requested coordinated orders only (including Number Portability orders where coordination is requested by the CLEC).
Notes:	• "Cutovers" include initial and subsequent attempts to complete a cutover. (Pacific Bell only)

Provisioning Measure 9A

Title: Frame Due Time Conversions as a Percentage On-Time - Pacific Bell only

Area	*Requirement Description*
Description:	Measures the percentage of Frame Due Time cutovers completed by Committed time* for all orders where CLEC has requested FDT. ** Note: "Committed time" means within 1 hour of confirmed frame due time (example: order with 4pm due time will be completed by 5pm).*
Method of Calculation:	(Number of frame due time cutovers completed by Committed time) / (Count of frame due time cutovers scheduled in reporting period)x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), by ILEC Affiliates
Reported By:	Basic loops with LNP, Basic loops without LNP, Standalone LNP.
Geographic Level:	Statewide
Measurable Standard:	**Benchmark** • Standard 95% in 1 hour
Business Rules:	• Excludes CLEC caused misses • Applies to CLEC requested FDT orders only
Notes:	• "Cutovers" include initial and subsequent attempts to complete a cutover. • Up to 19 loops, or up to 99 telephone numbers on standalone LNP.

OSS OII Performance Measurements
Report Requirements

Provisioning **Measure 10**

Title: LNP Network Provisioning

Area	*Requirement Description*
Description:	Measures LNP network provisioning failures as a percentage of the total number of NPAC broadcasts of telephone number subscription versions to port.
Method of Calculation:	(Total number of LNP network provisioning failures / Total number of NPAC porting broadcasts) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and ILEC Affiliates
Reported By:	
Geographic Level:	Statewide
Measurable Standard:	**Benchmark for Pacific Bell** • Standard - no more than .25% failure **Benchmark for GTE** • Standard - no more than 2% failure
Business Rules:	• Provisioning failure data will be collected as follows: • Will be tracked for individual network database failures - failures to provision between the ILEC LSMS and LNP network databases (STP or SCP) • Excludes total failures from the NPAC to *all* LSMS systems. • Excludes broadcasts failing due to a lack of GTT information made available to ILEC (no SS7 signaling agreement in place between ILEC and CLEC) (Pacific Bell only) • Excludes large porting activities (500 TNs or greater) (Pacific Bell only)
Notes:	

Provisioning **Measure 11**

Title: Percent of Due Dates Missed

Area	*Requirement Description*
Description:	Measures the percent of new, move and change orders where installation was not completed by the due date.
Method of Calculation:	[(Total Number of Missed Due Dates Due to ILEC Reasons for New, Move and Change Orders / Total Number of New, Move and Change Orders)] x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	By service group type and Field Work/No Field Work as appropriate
Geographic Level:	Region (PB), Statewide (GTE)

Measurable Standard:	**<u>Pacific Bell</u>** **Parity for Resale is Retail** **Parity for UNE measured for the following UNEs:**	**Pacific Bell Retail**
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX) • UNE Subloop	• POTS - Business (fielded)
	• 2w digital loop(ISDN capable) • UNE Subloop	• ISDN(BRI)
	• 2w digital loop(xDSL capable) • UNE Subloop	• 2w digital loop (xDSL capable) provided to ASI
	• 2w digital loop(IDSL capable) • UNE Subloop	• ISDN(BRI)

	• High Bandwidth line sharing UNE
	    • Conditioned
	    • Non-Conditioned
		• High Bandwidth line sharing UNE provided to ASI

	• 4w digital loop(DS1)	• DS1
	• UNE loop – DS3	• UNE loop – DS3
	• UNE loop – OC level service	• Retail OC level service
	• UNE Port–Non-Specials	• POTS - Business (non-fielded)
	• UNE Port– Specials	• Retail Specials (non-fielded)

	• UNE Dedicated Transport
	    • DS1
	    • DS3
	    • OC level
		• HICAP
		    • DS1
		    • DS3
		    • Retail OC level service

	• Dark Fiber	*Diagnostic*

	• Enhanced Extended Links
	    • VG - Conversion
	    • DS1 - New
	    • DS1 -Conversion
	    • DS3- New
	    • DS3-Conversion
	    • OC level - New
	    • OC level - Conversion
		(TBD)

	• UNE Platform
	    • Basic port and loop
	    • Special port and basic loop
	    • ISDN BRI port and loop
	    • ISDN PRI port and loop
		• Business POTS FW/NFW
		• Retail Voice Grade Specials FW/NFW
		• ISDN BRI FW/NFW
		• ISDN PRI FW/NFW

	• Interconnection Trunks	• ILEC Dedicated Trunks

Measurable Standard:	**GTE**	**Retail**
	• Resale POTS- Residence	• Retail POTS - Residence
	• Resale POTS-Business	• Retail POTS - Business
	• Resale Specials	• Retail Specials
	• UNE loop Nondesigned	• B1 Dispatched Non Designed
	• UNE loop Designed	• Dispatched Designed Service (excludes HICAPs)
	• UNE loop xDSL capable	• *(TBD until SDA is established)*
	• UNE loop IDSL capable	• *(TBD until SDA is established)*
	• UNE Port	• CentraNet - Simple
	• UNE Transport	• HICAP Designed
	• UNE Platform	
	• UNE - P Res	• Residential POTS
	• UNE - P Bus	• Business POTS
	• UNE - P PRI	• ISDN PRI
	• Interconnection Trunks	• ILEC Dedicated Trunks
	• Line Sharing - Conditioned	• *(TBD until SDA is established)*
	• Line Sharing - Non-Conditioned	• *(TBD until SDA is established)*
	• LNP	• Retail POTS - Total Business & Residence, Non-Dispatched
	• EEL	• *(Diagnostic)*
	• Subloop	• *(Diagnostic)*
	• Dark Fiber	• *(Diagnostic)*

Business Rules:	• Excludes customer misses • Due date is defined as either original due date or final due date if the original due date was missed due to customer reasons. • For UNE loop services, feature-only orders are excluded from retail analog. (Pacific Bell only) • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity ASI (Pacific Bell only) • For GTE results for UNE subloop will be tracked diagnostically. • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review. • Excludes record only and ILEC official orders. • The Completion Date is the date on which the service has passed acceptance testing, where applicable. To the extent that Pacific is required to obtain affirmative acceptance of the loop from the CLEC before closing an order, the order will not be deemed to have successfully passed an acceptance test until the CLEC affirmatively accepts the loop. (Pacific Bell only) • Orders where acceptance testing is delayed as a result of CLEC action or inaction shall be excluded. (Pacific Bell only)
Notes:	• ILECs will provide disaggregation by Missed Appointment reason codes as diagnostic data upon raw data request. • For Pacific Bell, no retail analog exists for IDSL capable loops. The retail comparison will be made with ISDN service which has similar characteristics

Provisioning — Measure 12

Title: Percent of Due Dates Missed Due to Lack of Facilities

Area	*Requirement Description*
Description:	Measures the percent of new, move and change orders missed due to lack of facilities. Note: Results also included in Measure "Percent Missed Due Dates"
Method of Calculation:	(Total New, Move and Change Orders Missed Due Dates Due to Lack of Facilities) / (Total Number of New, Move and Change Orders) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	By service group type and Field Work/No Field Work as appropriate
Geographic Level:	Region (PB), Statewide (GTE)

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail** **Parity measured for the following UNEs:**	**Retail**
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX)	• POTS - Business (fielded)
	• 2w digital loop(ISDN capable)	• ISDN(BRI)
	• 2w digital loop(xDSL capable)	• 2w digital loop(xDSL capable) provided to ASI
	• 2w digital loop(IDSL capable)	• ISDN (BRI)

	• High Bandwidth line sharing UNE
	  • Conditioned
	  • Non-Conditioned
		• High Bandwidth line sharing UNE provided to ASI

	• 4w digital loop (DS1)	• DS1
	• UNE loop – DS3	• DS3
	• UNE loop – OC level	• Retail OC level service

	• UNE Dedicated Transport
	  • DS1
	  • DS3
	  • OC level
		• HICAP
		  • DS1
		  • DS3
		  • Retail OC level service

	• Enhanced Extended Links
	  • DS1 - New
	  • DS3 – New
	  • OC level - New
		(TBD)

	• UNE Platform
	  • Basic port and loop
	  • Special port and basic loop
	  • ISDN BRI port and loop
	  • ISDN PRI port and loop
		
		• Business POTS FW/NFW
		• Retail Voice Grade Specials FW/NFW
		• ISDN BRI FW/NFW
		• ISDN PRI FW/NFW

	• Interconnection Trunks	• ILEC Dedicated Trunks

Measurable Standard:	**GTE**	**Retail**
	• Resale POTS- Residence	• Retail POTS - Residence
	• Resale POTS-Business	• Retail POTS - Business
	• Resale Specials	• Retail Specials
	• UNE loop Nondesigned	• B1 Dispatched Non Designed
	• UNE loop Designed	• Dispatched Designed Service (excludes HICAPs)
	• UNE loop xDSL capable	• *(TBD until SDA is established)*
	• UNE loop IDSL capable	• *(TBD until SDA is established)*
	• Line Sharing - Conditioned	• *(TBD until SDA is established)*
	• Line Sharing - Non-Conditioned	• *(TBD until SDA is established)*
	• UNE Port	• CentraNet - Simple
	• UNE Transport	• HICAP Designed
	• UNE Platform	
	• UNE - P Res	• Residential POTS
	• UNE - P Bus	• Business POTS
	• UNE - P PRI	• ISDN PRI
	• Interconnection Trunks	• ILEC Dedicated Trunks
	• EEL	• *(Diagnostic)*
	• Subloop	• *(Diagnostic)*
Business Rules:	• Due date is defined as either original due date or final due date if the original due date was missed due to customer reasons. • For UNE loop services, feature-only orders are excluded from retail analog.	
Notes:	• For Pacific Bell, no retail analog exists for IDSL capable loops. The retail comparison will be made with ISDN capable loops which have similar characteristics.	

Provisioning **Measure 13**

Title: Delay Order Interval to Completion Date (For Lack of Facilities)

Area	Requirement Description
Description:	Measures the average calendar days from due date to completion date on company missed orders due to lack of ILEC facilities.
Method of Calculation:	Sum (Completion Date - Committed Order Due Date (for orders missed due to lack of ILEC facilities)) / (Number of Orders Missed due to Lack of ILEC Facilities in the Reporting Period)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	• By service group type • Disaggregated by 1-30 days, 31-90 days and >90 days
Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail** **Parity measured for the following UNEs:**	**Retail**
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX)	• POTS - Business (fielded)
	• 2w digital loop(ISDN capable)	• ISDN(BRI)
	• 2w digital loop(xDSL capable)	• 2w digital loop (xDSL capable) provided to ASI
	• 2w digital loop (IDSL capable)	• ISDN(BRI)

	• High Bandwidth line sharing UNE
	  • Condition
	  • Non-Condition
		• High Bandwidth line sharing UNE provided to ASI

	• 4w digital loop (DS1)	• DS1
	• UNE loop – DS3	• DS3
	• UNE loop – OC level	• Retail OC level service

	• UNE Dedicated Transport
	  • DS1
	  • DS3
	  • OC level
		• HICAP
		  • DS1
		  • DS3
		  • Retail OC level service

	• Enhanced Extended Links
	  • DS1 - New
	  • DS3 – New
	  • OC level - New
		(TBD)

	• UNE Platform
	  • Basic port and loop
	  • Special port and basic loop
	  • ISDN BRI port and loop
	  • ISDN PRI port and loop
		• Business POTS FW/NFW
		• Retail Voice Grade Specials FW/NFW
		• ISDN BRI FW/NFW
		• ISDN PRI FW/NFW

	• Interconnection Trunks	• ILEC Dedicated Trunks

Measurable Standard:	**GTE**
	
	• Resale POTS- Residence
	• Resale POTS-Business
	• Resale Specials
	• UNE loop Nondesigned
	• UNE loop Designed
	• UNE loop xDSL capable
	• UNE loop IDSL capable
	• Line Sharing - Conditioned
	• Line Sharing - Non-Conditioned
	• UNE Port
	• UNE Transport
	• UNE Platform
	    • UNE - P Res
	    • UNE - P Bus
	    • UNE - P PRI
	• Interconnection Trunks
	• EEL
	• Subloop
		Retail
		
		• Retail POTS - Residence
		• Retail POTS - Business
		• Retail Specials
		• B1 Dispatched Non Designed
		• Dispatched Designed Service (excludes HICAPs)
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• CentraNet-Simple
		• HICAP Designed
		
		• Residential POTS
		• Business POTS
		• ISDN PRI
		• ILEC Dedicated Trunks
		• *(Diagnostic)*
		• *(Diagnostic)*

Business Rules:	• For UNE loop services, feature-only orders are excluded from retail analog.	
Notes:	• For Pacific Bell, no retail analog exists for IDSL capable loops. The retail comparison will be made with ISDN service which has similar characteristics.	

Provisioning **Measure 14**

Title: Held Order Interval

Area	*Requirement Description*
Description:	Measures the time period that service orders are not completed by the original due dates for all ILEC reasons (including lack of facilities).
Method of Calculation:	Sum (Reporting Period Close Date - Committed Order Due Date) / (Number of Orders Pending and Past the Committed Due Date) *Note: For all orders pending and past the committed due date.*
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), by ILEC Affiliates
Reported By:	By service group type
Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail** **Parity for UNE measured for the following UNEs:**	**Retail**
	• 2/4w (8db and 5.5 db) analog loop (incl. Coin/analog PBX) • UNE Subloop	• POTS - Business (fielded)
	• 2w digital loop(ISDN capable) • UNE Subloop	• ISDN(BRI)
	• 2w digital loop(xDSL capable) • UNE Subloop	• 2w digital loop(xDSL capable) provided to ASI
	• 2w digital loop (IDSL capable) • UNE Subloop	• ISDN(BRI)

	• High Bandwidth line sharing UNE
	  • Conditioned
	  • Non-Conditioned
		• High Bandwidth line sharing UNE provided to ASI

	• 4w digital loop (DS1) • UNE Subloop	• DS1
	• UNE loop – DS3	• DS3
	• UNE loop – OC level	• Retail OC level service
	• UNE Port–Non-Specials	• POTS - Business (non-fielded)
	• UNE Port– Specials	• Retail Specials

	• UNE Dedicated Transport
	  • DS1
	  • DS3
	  • OC Level
		• HICAP
		  • DS1
		  • DS3
		  • Retail OC level service

	• Dark Fiber	• Diagnostic

	• Enhanced Extended Links
	  • VG - Conversion
	  • DS1 - New
	  • DS1 -Conversion
	  • DS3- New
	  • DS3-Conversion
	  • OC level – New
	  • OC level - Conversion
		(TBD)

	• UNE Platform (PB only)
	  • Basic port and loop
	  • Special port and basic loop
	  • ISDN BRI port and loop
	  • ISDN PRI port and loop
		• Business POTS FW/NFW
		• Retail Voice Grade Specials FW/NFW
		• ISDN BRI FW/NFW
		• ISDN PRI FW/NFW

	• Interconnection Trunks	• ILEC Dedicated Trunks

Measurable Standard:	**GTE**	**Retail**
	• Resale POTS- Residence	• Retail POTS - Residence
	• Resale POTS-Business	• Retail POTS - Business
	• Resale Specials	• Retail Specials
	• UNE loop Nondesigned	• B1 Dispatched Non Designed
	• UNE loop Designed	• Dispatched Designed Service (excludes HICAPs)
	• UNE loop xDSL capable	• *(TBD until SDA is established)*
	• UNE loop IDSL capable	• *(TBD until SDA is established)*
	• UNE Port	• CentraNet-Simple
	• UNE Transport	• HICAP Designed
	• UNE Platform	
	• UNE - P Res	• Residential POTS
	• UNE - P Bus	• Business POTS
	• UNE - P PRI	• ISDN PRI
	• Interconnection Trunks	• ILEC Dedicated Trunks
	• Line Sharing - Conditioned	• *(TBD until SDA is established)*
	• Line Sharing - Non-Conditioned	• *(TBD until SDA is established)*
	• LNP	• Retail POTS - Total Business & Residence, Non-Dispatched
	• EEL	• *(Diagnostic)*
	• Subloop	• *(Diagnostic)*
	• Dark Fiber	• *(Diagnostic)*
Business Rules:	• Excludes customer caused misses. • For UNE loop services, feature-only orders are excluded from retail analog. • The Completion Date is the date on which the service has passed acceptance testing, where applicable. To the extent that Pacific is required to obtain affirmative acceptance of the loop from the CLEC before closing an order, the order will not be deemed to have successfully passed an acceptance test until the CLEC affirmatively accepts the loop. (Pacific Bell only) • Orders where acceptance testing is delayed as a result of CLEC action or inaction shall be excluded. (Pacific Bell only)	

Notes:	• ILECs will provide disaggregation by Missed Appointment reason codes as diagnostic data upon raw data request. • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity ASI (Pacific Bell only) • For GTE results for UNE subloop will be tracked diagnostically. • For Pacific Bell, no retail analog exists for IDSL capable loops. The retail comparison will be made with ISDN capable loops which have similar characteristics.

Provisioning **Measure 15**

Title: Provisioning Trouble Reports (Prior to Service Order Completion)

Area	*Requirement Description*
Description:	Measures the percent of troubles that are reported (via customer or indirectly by CLEC) that occur during the provisioning process.
Method of Calculation:	**Parity:** (Number of trouble reports that occur from the time of service order creation, up to and including the date of service order completion)/ (Total Number of service orders in reporting period) **Benchmark:** [(Number of trouble reports that occur from the time of service order creation, up to and including the date of service order completion)/ (Total Number of service orders in reporting period)] x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), by ILEC Affiliates
Reported By:	• By Resale, High Bandwidth line sharing UNE, UNE Loop, and LNP • By Affecting Service and Out of Service
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell:** **Parity** **Resale** — **Retail services** UNE Loop — Retail services (outside plant disposition codes and central office wiring disposition codes) High Bandwidth Line sharing UNE — High Bandwidth line sharing UNE provided to ASI **Benchmark:** LNP - Port Out • Standard - 1% or less

	GTE: • Resale POTS (Residence) • Resale POTS (Business) • Resale Specials • UNE,Loop Non-designed • UNE Loop Designed • UNE Loop xDSL Capable • UNE Loop IDSL Capable • LNP	 • Residence POTS • Business POTS • Retail Specials • B1 Dispatched Non Designed • Dispatched Designed Service (excludes HICAPs) • *(TBD until SDA is established)* • *(TBD until SDA is established)* • *(TBD- will propose benchmark standard after 4 months of data collection).*
Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes Subsequent reports • Excludes Message Reports (circuit reports for which ILEC has no records) • Excludes ILEC employee generated reports • *[6]	
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request.	

[6] The language "excludes new service installations" first contained in the JPSA filed July 18, 2000 has been removed pending resolution by the Commission of the open issue identified by some DSL CLECs.

OSS OII Performance Measurements
Report Requirements

Provisioning **Measure 15A**

Title: Average Time to Restore Provisioning Troubles (Prior to Service Order Completion)

Area	*Requirement Description*
Description:	Measures the average duration of the troubles from the receipt of the customer trouble reported (via customer or indirectly by CLEC) to the time the trouble is cleared.
Method of Calculation:	(Total duration of provisioning trouble measured from the time the trouble was initiated or called in to the ILEC until cleared.)/ (Total Number of Provisioning Trouble Reports)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), by ILEC Affiliates
Reported By:	• By Resale, UNE Loop, UNE Port and LNP • By Affecting Service and Out of Service
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell:** **Parity:** **Resale** — **Retail services** UNE Loop — Retail services (outside plant disposition codes and Central Office wiring disposition codes) **Benchmark:** LNP - Port Out • Standard - average of 4 hours

<table>
<tr><td>Measurable Standard:</td><td>GTE
Resale POTS- Residence
Resale POTS-Business
Resale Specials
UNE loop Nondesigned
UNE loop Designed
UNE loop xDSL capable
UNE loop IDSL capable
LNP</td><td>Retail
Residence POTS
Business POTS
Retail Specials
B1 Dispatched Non Designed
Dispatched Designed Service (excludes HICAPs)
(TBD until SDA is implemented)
(TBD until SDA is implemented)
(TBD)</td></tr>
<tr><td>Business Rules:</td><td colspan="2">Excludes CPE and IEC/CLEC caused troubles
Excludes Subsequent reports
Excludes Message Reports (circuit reports for which ILEC has no records)
Excludes ILEC employee generated reports</td></tr>
<tr><td>Notes:</td><td colspan="2">ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request.</td></tr>
</table>

OSS OII Performance Measurements
Report Requirements

<u>Provisioning</u> **Measure 16**

Title: Percentage Troubles in 30 Days for Special Services Orders

Area	*Requirement Description*
Description:	Measures the percent of network customer trouble reports received within 30 calendar days of service order completion
Method of Calculation:	**Pacific Bell:** (Total Number of Customer Trouble reports received within 30 calendar days of special service order completion / Total Number of new, move and change completed special services orders) x 100 **GTE:** (Total Number of Special Service Orders that receive a Network Customer Trouble Report within 30 calendar days of service order completion / Total new, move and change completed Special Service orders) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	By service group type
Geographic Level:	Region (PB), Statewide (GTE)

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail**	
	Parity for UNE measured for the following UNEs:	**Retail**
	• 2w digital loop(ISDN capable) • UNE Sub -Loop	• ISDN(BRI) (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop(xDSL capable) • UNE Sub-Loop	• 2w digital loop(xDSL capable) provided to ASI (outside plant disposition codes and central office wiring disposition codes)
	• High Bandwidth line sharing UNE	• High Bandwidth line sharing UNE provided to ASI
	• 4w digital loop (DS1)	• DS1 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop – DS3	• DS3 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop –OC level	• Retail OC level service (outside plant disposition codes and central office wiring disposition codes)
	• UNE Port– Specials	• Retail Special (non-dispatched)

	• UNE Dedicated Transport
	  • DS1
	  • DS3
	  • OC level
		• HICAP
		  • DS1
		  • DS3
		  • Retail OC level

	• Dark Fiber	Diagnostic

	• Enhanced Extended Links
	  • VG - Conversion
	  • DS1 - New
	  • DS1 -Conversion
	  • DS3- New
	  • DS3-Conversion
	  • OC level – New
	  • OC level - Conversion
		(TBD)

	• UNE Platform
	  • Special port and basic loop
	  • ISDN BRI port and loop
	  • ISDN PRI port and loop
		• Retail Voice Grade Specials (non-disp, disp)
		• ISDN BRI (non-disp, disp)
		• ISDN PRI (non-disp, disp)

	• Interconnection Trunks	• ILEC Dedicated Trunks

Measurable Standard:	**GTE:** • Resale Specials • UNE Loop Designed • UNE loop xDSL capable • UNE loop IDSL capable • UNE Transport • UNE - Platform PRI • Line Sharing – Conditioned • Line Sharing - Non - Conditioned • Interconnection Trunks • EEL	**Retail** • Retail Specials • Dispatch Designed Service (excludes HICAPs) • (TBD until SDA is established) • (TBD until SDA is established) • HICAP Designed • ISDN PRI • (TBD until SDA is established) • (TBD until SDA is established) • ILEC Dedicated Trunks • *(Diagnostic)*
Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes troubles associated with inside wire • Excludes Trouble Reports Received on the Due Date (which instead are reported in the "Provisioning Troubles" measure) • Excludes Subsequent reports • Excludes Message Reports (circuit reports for which ILEC has no records) • Excludes ILEC employee generated reports • If no service orders are processed for a service group type in the report month, the denominator for the calculation of this measure will be service orders processed in the last month of service order activity. (Pacific Bell) • The Completion Date is the date on which the service has passed acceptance testing, where applicable. To the extent that Pacific is required to obtain affirmative acceptance of the loop from the CLEC before closing an order, the order will not be deemed to have successfully passed an acceptance test until the CLEC affirmatively accepts the loop. (Pacific Bell only) • Orders where acceptance testing is delayed as a result of CLEC action or inaction shall be excluded. (Pacific Bell only)	
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity ASI (Pacific Bell only) • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	

OSS OII Performance Measurements
Report Requirements

Provisioning — **Measure 17**

Title: Percentage Troubles in 7 Days for Non-Special Orders - GTE only
Percentage Trouble in 10 Days for Non-Special Orders - Pacific Bell only

Area	*Requirement Description*
Description:	Measures the percent of network customer trouble reports received within 7 (GTE) or 10 (Pacific Bell) calendar days of service order completion.
Method of Calculation:	**GTE:** (Total Number of non-special Service Orders that receive a Network Customer Trouble Report within 7 calendar days of service order completion / Total new, move and change completed Non-Special Service orders) x 100 **Pacific Bell:** (Total Number of Customer Trouble reports received within 10 calendar days of non-special service order completion / Total Number of new, move and change completed non-special orders) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	By service group type (including LNP) and Field Work/No Field Work as appropriate
Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail (non-special services only)** **Parity for UNE measured for the following UNEs:** • 2/4w (8db and 5.5 db) loop (incl. Coin/analog PBX) 　• UNE Sub-Loop (and for Pacific Bell only) 　• FDT orders 　• TBCC orders • UNE Port – Basic analog/Coin • UNE Platform -Basic port and basic loop • LNP (Port Out)	 **Retail** • Business POTS (outside plant disposition codes and central office wiring disposition codes) • Business POTS (non-disp) • Business POTS (disp/non-disp) • Benchmark of no more than 1% troubles.
	GTE • Resale POTS- Residence • Resale POTS-Business • UNE loop Nondesigned • • UNE Port • UNE Platform 　• UNE - P Res 　• UNE - P Bus • LNP • Subloop	**Retail** • Retail POTS - Residence • Retail POTS - Business • B1 Dispatched Non Designed • CentraNet - Simple • Residential POTS • Business POTS • Retail POTS- Total Business & Residence, Non-Dispatched • *(Diagnostic)*

Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes Trouble Reports Received on the Due Date • Excludes Subsequent reports • Excludes ILEC employee generated reports • Excludes troubles associated with inside wiring. • If no service orders are processed for a service group type in the report month, the denominator for the calculation of this measure will be service orders processed in the last month of service order activity. (Pacific Bell only) • The Completion Date is the date on which the service has passed acceptance testing, where applicable. To the extent that Pacific is required to obtain affirmative acceptance of the loop from the CLEC before closing an order, the order will not be deemed to have successfully passed an acceptance test until the CLEC affirmatively accepts the loop. (Pacific Bell only) • Orders where acceptance testing is delayed as a result of CLEC action or inaction shall be excluded. (Pacific Bell only)
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type. • Pacific Bell will track FDT and TBCC diagnostically until the next review cycle.

OSS OII Performance Measurements
Report Requirements

Provisioning **Measure 18**

Title: Completion Notice Interval

Area	*Requirement Description*
Description:	Measures the percent of completion notices returned within the time specified in the measurable standard.
Method of Calculation:	**Fully Electronic:** (Number of Completion Notices Returned within “X” Interval) / (Number of Orders Completed where the Completion Notice is Returned Using Electronic Process) x 100 **All Other Interfaces:** (Number of Completion Notices Returned within “X” Interval) / (Number of Orders Returned Using All Other Processes) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, and by ILEC Affiliates
Reported By:	All interfaces
Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell:** **Fully electronic(LEX, EDI)** - • Standard -95% within 1hour **Fully electronic Fallout**: • Standard is 95% within 24 hours with a fallout maximum of 5% for each system reported. If LASR shows a reduction in fallout level (an average to nearest 0.5%) for three reported months, then Pacific Bell will lower fallout level to match. **All other interfaces** • Standard– 90% within 24 hours **GTE:** **Fully Electronic (EDI)** • Standard - 95% within 1 hour Electronic Batch • Standard – 95% within 12 hours **All other interfaces** • Standard – 90% within 24 hours
Business Rules:	• 24 hour clock is used to measure interval for all other interfaces. • Excludes weekends and ILEC published holidays • System hours will be used for fully electronic sub-measures • GTE will report on the industry standard of SAR Version 4 only. • For GTE, fully electronic represents all near "real-time" interfaces that flow through and do not include batch processing. • For GTE, Electronic Batch represents all electronic interfaces that include some form of batch processing. • For GTE, all other interfaces represent manual processes. • For GTE, Electronic Batch will use the same calculation method as Fully Electronic
Notes:	• Completion Notices on disconnect orders are only for CLEC disconnect orders (not on ILEC retail disconnect orders, except for LNP disconnect orders).

Maintenance — Measure 19

Title: Customer Trouble Report Rate

Area	*Requirement Description*
Description:	Measures the total number of network customer trouble reports received within a calendar month per 100 local exchange lines/interconnection or interoffice trunks/ circuits/UNEs.
Method of Calculation:	(Total Number of Customer initial and repeat network trouble reports / Number of local exchange lines/interconnection or interoffice trunks/circuits/UNEs in service at the end of the prior reporting period) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Report By:	By service group type (including LNP) & NXX Code Opening Troubles
Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail** **Parity for UNE measured for the following UNEs:**	**Retail**
	• 2/4w (8db and 5.5db) analog loop	• POTS - Business (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop (ISDN)	• ISDN(BRI) (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop (xDSL)	• 2w digital loop (xDSL) provided to ASI (outside plant disposition codes and central office wiring disposition codes)
	• High Bandwidth line sharing UNE	• High Bandwidth line sharing UNE provided to ASI
	• 4w digital loop (DS1)	• DS1(outside plant disposition codes and central office wiring disposition codes)
	• UNE loop – DS3	• DS3 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop – OC level	• Retail OC level service (outside plant disposition codes and central office wiring disposition codes)
	• UNE Port – Non-Specials	• POTS - Business (dispatch in)
	• UNE Port – Specials	• Retail Specials (dispatch in)

	• UNE Dedicated Transport
	  • DS1
	  • DS3
	  • OC level
		• HICAP
		  • DS1
		  • DS3
		  • Retail OC level service

	• Dark Fiber	Diagnostic

	• Enhanced Extended Links
	  • VG
	  • DS1
	  • DS3
	  • OC level
		(TBD)

	• UNE Platform
	  • Basic port and loop
	  • Special port and basic loop
	  • ISDN BRI port and loop
	  • ISDN PRI port and loop
		
		• Business POTS (non-disp, disp)
		• Retail Voice Grade Specials (non-disp, disp)
		• ISDN BRI (non-disp, disp)
		• ISDN PRI (non-disp, disp)

	• Interconnection Trunks	• ILEC Dedicated Trunks
	• LNP - Port Out	• Benchmark: .35%

Measurable Standard:	**<u>GTE</u>**
	• Resale POTS- Residence
	• Resale POTS-Business
	• Resale Specials
	• UNE loop Nondesigned
	• UNE loop Designed
	• UNE loop xDSL capable
	• UNE loop IDSL capable
	• UNE Port
	• UNE Transport
	• UNE Platform
	  • UNE - P Res
	  • UNE - P Bus
	  • UNE - P PRI
	• Interconnection Trunks
	• Line Sharing - Conditioned
	• Line Sharing - Non - Conditioned
	• LNP
	
	• EEL
	• Dark Fiber
	• UNE Subloop
		Retail
		• Retail POTS - Residence
		• Retail POTS - Business
		• Retail Specials
		• B1 Dispatched Non Designed
		• Dispatched Designed Service (excludes HICAPs)
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• CentraNet-Simple
		• HICAP Designed
		
		• Residential POTS
		• Business POTS
		• ISDN PRI
		• ILEC Dedicated Trunks
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• No more than .35% of total trouble reports received for LNP
		• *(Diagnostic)*
		• *(Diagnostic)*
		• *(Diagnostic)*

Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes Subsequent reports • Excludes Message Reports (circuit reports for which ILEC has no records) • Access line/circuit count taken from previous month • Excludes ILEC employee generated reports • For GTE - excludes provisioning trouble reports. • Include Test okay (TOK) and Found Okay (FOK) reports.	
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type. (GTE only) • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	

Maintenance **Measure 20**

Title: Percentage of Customer Trouble Not Resolved Within Estimated Time

Area	*Requirement Description*
Description:	Measures the percent of trouble reports not cleared by the commitment time.
Method of Calculation:	(Total network trouble reports not cleared by the commitment time for ILEC reasons / Total network trouble reports completed) x 100
Report Period:	Monthly
Report Structure :	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Report By:	• By service group type (including LNP) & NXX Code Opening Troubles • By dispatch and no dispatch
Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell** **Parity for Resale is Retail**	
	Parity for UNE measured the following UNEs: 2/4w (8db and 5.5db) analog loop • UNE Sub-Loop	**Retail** • POTS - Business (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop (ISDN) • UNE Sub-Loop	• ISDN(BRI) (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop (xDSL) • UNE Sub-Loop	• 2w digital loop (xDSL) provided to ASI (outside plant disposition codes and central office wiring disposition codes)
	• High Bandwidth line sharing UNE	• High Bandwidth line sharing UNE provided to ASI
	• 4w digital loop (DS1) • UNE Subloop	• DS1 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop –DS3	• DS1 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop – OC level	• Retail OC level service (outside plant disposition codes and central office wiring disposition codes)
	• UNE Port – Non Specials	• POTS - Business (dispatch in)
	• UNE Port – Specials	• Retail Specials(dispatch in)

	• UNE Dedicated Transport
	    • DS1
	    • DS3
	    • OC level
		• HICAP
		    • DS1
		    • DS3
		    • Retail OC level service

	• Dark Fiber	Diagnostic

	• Enhanced Extended Links
	    • VG
	    • DS1
	    • DS3
	    • OC level
		(TBD)

	• UNE Platform
	    • Basic port and loop
	    • Special port and basic loop
	    • ISDN BRI port and loop
	    • ISDN PRI port and loop
		• Business POTS non-disp,disp)
		• Retail Voice Grade Specials (non-disp, disp)
		• ISDN BRI (non-disp, disp)
		• ISDN PRI (non-disp,disp)

	• Interconnection Trunks	• ILEC Dedicated Trunks
	• LNP - Port Out	• Benchmark**:** No more than 1 missed commit per month per CLEC

Measurable Standard:	**GTE**	**Retail**
	• Resale POTS- Residence	• Retail POTS - Residence
	• Resale POTS-Business	• Retail POTS - Business)
	• Resale Specials	• Retail Specials
	• UNE loop Nondesigned	• B1 Dispatched Non Designed
	• UNE loop Designed	• Dispatched Designed Service (excludes HICAPs)
	• UNE loop xDSL capable	• *(TBD until SDA is established)*
	• UNE loop IDSL capable	• *(TBD until SDA is established)*
	• UNE Port	• CentraNet - Simple
	• UNE Transport	• HICAP Designed
	• UNE Platform	
	◦ UNE - P Res	• Residential POTS
	◦ UNE - P Bus	• Business POTS
	◦ UNE - P PRI	• ISDN PRI
	• Interconnection Trunks	• ILEC Dedicated Trunks
	• Line Sharing - Conditioned	• *(TBD until SDA is established)*
	• Line Sharing - Non - Conditioned	• *(TBD until SDA is established)*
	• LNP	• No more than 1 missed commit per month per CLEC
	• EEL	• *(Diagnostic)*
	• Dark Fiber	• *(Diagnostic)*
	• UNE Subloop	• *(Diagnostic)*
Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes Subsequent reports • Excludes Message Reports (circuit reports which ILEC has no records on) • Excludes ILEC employee generated reports • Excludes customer caused misses • Results include Test okay (TOK) and Found Okay (FOK) reports. • For GTE - excludes provisioning trouble reports.	

Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity ASI (Pacific Bell only) • Results for UNE Subloops will be tracked diagnostically (GTE only) • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.

Maintenance **Measure 21**

Title: Average Time to Restore

Area	*Requirement Description*
Description:	Measures the average duration of customer trouble reports from the receipt of the customer trouble report to the time the trouble is cleared.
Method of Calculation:	(Total duration of customer network trouble reports) / (Total customer network trouble reports)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	• By service group type (including LNP) & NXX Code Opening Troubles • By dispatch and no dispatch
Geographic Level:	Statewide

Measurable Standard:

Pacific Bell
Parity for Resale is Retail

Parity for UNE measured for the following UNEs:	**Retail**
• 2/4w (8db and 5.5 db) analog loop • UNE Sub-Loop	• POTS - Business (outside plant disposition codes and central office wiring disposition codes)
• 2w digital loop (ISDN) • UNE Sub-Loop	• ISDN(BRI) (outside plant disposition codes and central office wiring disposition codes)
• 2w digital loop (xDSL) • UNE Sub-Loop	• 2w digital loop (xDSL) provided to ASI (outside plant disposition codes and central office wiring disposition codes)
• High Bandwidth line sharing UNE	• High Bandwidth line sharing UNE provided to ASI
• 4w digital loop (DS1) • UNE Sub-Loop	• DS1 (outside plant disposition codes and central office wiring disposition codes)
• UNE Loop – DS3	• DS3 (outside plant disposition codes and central office wiring disposition codes)
• UNE loop – OC level	• Retail OC level service (outside plant disposition codes and central office wiring disposition codes)
• UNE Port – Non-Specials	• POTS - Business (dispatch in)
• UNE Port – Specials	• Retail Specials (dispatch in)

• UNE Dedicated Transport
  • DS1
  • DS3
  • OC level
	• HICAP
	  • DS1
	  • DS3
	  • Retail OC level service

• Dark Fiber	• Diagnostic

• Enhanced Extended Links
  • VG
  • DS1
  • DS3
  • OC level
	(TBD)

• UNE Platform
  • Basic port and loop
  • Special port and basic loop
  • ISDN BRI port and loop
  • ISDN PRI port and loop
	
	• Business POTS (non-disp, disp)
	• Retail Voice Grade Specials (non-disp, disp)
	• ISDN BRI (non-disp, disp)
	• ISDN PRI (non-disp, disp)

• Interconnection Trunks	• ILEC Dedicated Trunks
• LNP - Port Out	• Benchmark: avg. 4 hours

Measurable Standard:	GTE
	
	• Resale POTS- Residence
	• Resale POTS-Business
	• Resale Specials
	• UNE loop Nondesigned
	• UNE loop Designed
	• UNE loop xDSL capable
	• UNE loop IDSL capable
	• UNE Port
	• UNE Transport
	• UNE Platform
	    • UNE - P Res
	    • UNE - P Bus
	    • UNE - P PRI
	• Interconnection Trunks
	• Line Sharing - Conditioned
	• Line Sharing - Non - Conditioned
	• LNP
	
	• EEL
	• Dark Fiber
	• UNE Subloop
		Retail
		
		• Retail POTS - Residence
		• Retail POTS - Business
		• Retail Specials
		• B1 Dispatched Non Designed
		• Dispatched Designed Service (excludes HICAPs)
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• CentraNet - Simple
		• HICAP Designed
		
		• Residential POTS
		• Business POTS
		• ISDN PRI
		• ILEC Dedicated Trunks
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• *Retail POTS – Total Business & Residence, Non-Dispatched*
		• *(Diagnostic)*
		• *(Diagnostic)*
		• *(Diagnostic)*

Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes Subsequent reports • Excludes Message Reports (circuit reports which ILEC has no records on) • Excludes ILEC employee generated reports • For GTE - excludes provisioning trouble reports. • Results include Test okay (TOK) and Found Okay (FOK) reports.	
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request • Results for UNE Subloops will be tracked diagnostically, by UNE loop type except for xDSL subloop the measurable standard for which will be parity ASI (Pacific Bell only) • Results for UNE Subloops will be tracked diagnostically (GTE only) • Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	

OSS OII Performance Measurements
Report Requirements

Maintenance **Measure 22**

Title: POTS Out of Service Less Than 24 Hours

Area	*Requirement Description*
Description:	Measures the percent of POTS out-of-service trouble reports cleared in less than 24 hours.
Method of Calculation:	(Total number of out of service network troubles cleared in less than 24 hours / Total number of out of service network troubles reported) x 100 *Note: For non-design services only*
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Reported By:	By POTS Residence and Business (Resale and UNE)
Geographic Level:	Statewide
Measurable Standard:	**Parity for Resale (POTS) for Pacific Bell** **Parity for UNEs (Basic)** / **Retail** • 2/4w (8db and 5.5 db) analog loop • UNE Sub-Loop / • POTS - Business (dispatch) (outside plant disposition codes and central office wiring disposition codes) • UNE Port – Basic Analog / • POTS - Business (dispatch in) • UNE Platform – Basic Port and Loop / • Business POTS (non-disp/dispatch)_

	GTE / **Retail**
	• Resale POTS- Residence / • Retail POTS - Residence
	• Resale POTS-Business / • Retail POTS - Business
	• UNE loop Non-designed / • B1 Dispatched Non Designed
	• UNE Port / • CentraNet - Simple
	• UNE Platform
	  • UNE - P Res / • Residential POTS
	  • UNE - P Bus / • Business POTS

Business Rules:	• Residential and Business POTS only • Excludes no access • Interval for tickets received Saturday and Sunday begins no later than Monday morning • Excludes CPE and IEC/CLEC caused troubles • Excludes Subsequent reports • Excludes Message Reports (circuit reports for which ILEC has no records) • Excludes ILEC employee generated reports • Results include Test okay (TOK) and Found okay (FOK) reports.
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request. • Results for UNE Subloops will be tracked diagnostically, by UNE loop type (Pacific Bell only).

OSS OII Performance Measurements
Report Requirements

Maintenance **Measure 23**

Title: Frequency of Repeat Troubles in 30 Day Period

Area	*Requirement Description*
Description:	Measures the percent of customer network trouble reports received within 30 calendar days of a previous report.
Method of Calculation:	(Total customer network trouble reports received within 30 calendar days of a previous customer report / Total customer network trouble reports) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), and by ILEC Affiliates
Report By:	By service group type (including LNP) & NXX Code Opening Troubles
Geographic Level	Statewide

Measurable Standard:	**Pacific Bell**
	Parity for Resale is Retail
	
	Parity for UNE measured for the following UNEs: / **Retail**
	• 2/4w (8bd and 5.5db) analog loop — • POTS - Business (fielded) (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop (ISDN) — • ISDN(BRI) (outside plant disposition codes and central office wiring disposition codes)
	• 2w digital loop (xDSL) — • 2w digital loop (xDSL) provided to ASI (outside plant disposition codes and central office wiring disposition codes)
	• High Bandwidth line sharing UNE — • High Bandwidth line sharing UNE provided to ASI
	• 4w digital loop (DS1) — • DS1 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop – DS3 — • DS3 (outside plant disposition codes and central office wiring disposition codes)
	• UNE loop – OC level — • Retail OC level service (outside plant disposition codes and central office wiring disposition codes)
	• UNE Port – Non-Specials — • POTS - Business (dispatch in)
	• UNE Port –Specials — • Retail Specials (non-dispatch)
	• UNE Dedicated Transport (• DS1 • DS3 • OC level) — • HICAP (• DS1 • DS3 • Retail OC level service)
	• Dark Fiber — • Diagnostic
	• Enhanced Extended Links (• VG • DS1 • DS3 • OC level) — *(TBD)*
	• UNE Platform
	  • Basic port and loop — • Business POTS (non-disp, disp)
	  • Special port and basic loop — • Retail Voice Grade Specials (non-disp,disp)
	  • ISDN BRI port and loop — • ISDN BRI (non-disp, disp)
	  • ISDN PRI port and loop — • ISDN PRI (non-disp, disp)
	• Interconnection Trunks — • ILEC Dedicated Trunks
	• LNP - Port Out — • Benchmark**:** No more than 2 repeat troubles per month per CLEC

Measurable Standard:	**GTE**
	
	• Resale POTS- Residence
	• Resale POTS-Business
	• Resale Specials
	• UNE loop Nondesigned
	• UNE loop Designed
	• UNE loop xDSL capable
	• UNE loop IDSL capable
	• UNE Port
	• UNE Transport
	• UNE Platform
	    • UNE - P Res
	    • UNE - P Bus
	    • UNE - P PRI
	• Interconnection Trunks
	• Line Sharing - Conditioned
	• Line Sharing - Non - Conditioned
	• LNP
	
	• EEL
	• Dark Fiber
	• UNE Subloop
		Retail
		
		• Retail POTS - Residence
		• Retail POTS - Business
		• Retail Specials
		• B1 Dispatched Non Designed
		• Dispatched Designed Service (excludes HICAPs)
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• CentraNet - Simple
		• HICAP Designed
		
		• Residential POTS
		• Business POTS
		• ISDN PRI
		• ILEC Dedicated Trunks
		• *(TBD until SDA is established)*
		• *(TBD until SDA is established)*
		• No more than 2 repeat trouble per month per CLEC
		
		• *(Diagnostic)*
		• *(Diagnostic)*
		• *(Diagnostic)*

Business Rules:	• Excludes CPE and IEC/CLEC caused troubles • Excludes troubles associated with inside wiring • Excludes Subsequent reports • Excludes Message Reports • Excludes ILEC employee generated reports	
Notes:	• ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request.	

Network Performance **Measure 24**

Title: Percent Blocking on Common Trunks

Area	*Requirement Description*
Description:	Measures the percent of common and shared transport trunk groups exceeding 2% blockage.
Method of Calculation:	(Number of common and shared transport trunk groups exceeding 2% blockage / Total number of common and shared transport trunk groups) x 100
Report Period:	Monthly (Exception Reporting Only)
Report Structure:	
Report By:	By total trunk groups.
Geographic Level:	Statewide
Measurable Standard:	Benchmark: 2% of trunk groups blocking at no more than 2%
Business Rules:	• GTE reports provided 45 days after close of data month. • ILEC will make available detailed information for all trunk groups not meeting 2% blocking level with the monthly report
Notes:	

OSS OII Performance Measurements
Report Requirements

Network Performance **Measure 25**

Title: Percent Blocking on Interconnection Trunks

Area	*Requirement Description*
Description:	Measures the percent of final dedicated interconnection trunk groups exceeding 2% blockage. .
Method of Calculation:	(Number of final dedicated interconnection trunk groups exceeding 2% blockage / Total number of final dedicated interconnection trunk groups) x 100
Report Period:	Monthly (Exception Reporting Only)
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies), by ILEC Affiliates
Report By:	• Total trunk groups • ILEC end office to CLEC end office • ILEC tandem to CLEC end office
Geographic Level:	Statewide
Measurable Standard:	Parity for Pacific Bell and GTE – comparison made to ILEC final trunk groups
Business Rules:	• Only measured on trunks where ILEC has outgoing traffic to CLECs, and where ILEC controls trunk capacity. • GTE reports provided 45 days after close of data month. • Excludes blocking failures caused by the CLEC not completing growth trunk provisioning by scheduled due date. • Excludes blocking due to CLEC putting trunks in a "make busy" state. • Applies to those trunks where the ILEC has augmentation control. • Does not apply when trunks are provisioned as two-way trunks
Notes:	• ILEC will provide detail available regarding exclusions in raw data.

OSS OII Performance Measurements
Report Requirements

Network Performance — Measure 26

Title: NXX Loaded by LERG Effective Date

Area	Requirement Description
Description:	Measures the number of NXXs loaded and tested by the LERG effective date.
Method of Calculation:	((Number of NXXs loaded and tested by LERG effective date) / (Number of NXXs scheduled to be loaded and tested by LERG effective date)) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies)and by ILEC Affiliates
Report By:	Reported for all NXX codes scheduled to be loaded in reporting period
Geographic Level:	Statewide
Measurable Standard:	Parity for Pacific Bell and GTE – comparison made to results for loading ILEC NXX codes by the LERG effective date.
Business Rules:	• Excludes any NXX codes with requested loading interval of less than the industry standard (currently 45 days). • Excludes any NXX code that cannot be completely tested because the CLEC has not provided an accurate test number or because CLEC facilities have not been installed. • Includes both additions and deletions to NXX codes.
Notes:	• NXX loading procedures include central office/tandem translations, verification of translations, call through testing, and AMA testing. • TRUCALL billing validation testing is not used unless maintenance trouble is reported (Pacific Bell only)

Network Performance **Measure 27**

***Title*:** **MEASURE DELETED**

Area	*Requirement Description*
Description:	*Measure deleted - process is parity by design.*
Method of Calculation:	
Report Period:	
Report Structure:	
Report By:	
Geographic Level:	
Measurable Standard:	
Business Rules:	
Notes:	

OSS OII Performance Measurements
Report Requirements

Billing **Measure 28**

Title: Usage Timeliness

Area	*Requirement Description*
Description:	This measure captures the elapsed time between the recording of usage data generated either by CLEC retail customers or access usage associated with CLEC customers and the time when the data set, in a compliant format, is successfully transmitted to the CLEC.
Method of Calculation:	Sum ((Data Set Transmission Availability Date) - (Date of Message Recording)) / (Count of All Messages available for Transmission in Reporting Period)
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and by ILEC Affiliates
Report By:	**Pacific Bell:** • Resale • UNE (IntraLATA and InterLATA, combined) • Jointly provided switched access (associated with meet point billing) **GTE** • Resale Local • Resale Toll • UNE (IntraLATA and InterLATA combined)(excluding UNE Platform) • UNE Platform – Local • UNE Platform - Access • Jointly provided switched access (associated with meet point billing)
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell:** Parity for Resale UNE, and Jointly provided switched access: **GTE:** Parity for Resale - Local, Resale - Toll and UNE Parity for UNE Platform – Local is Resale – Local Parity for UNE Platform – Access is IXC switched access Benchmark for Jointly provided switched access: Standard – 95% in 6 Days
Business Rules:	
Notes:	• GTE bills local/toll through CBSS billing systems. Access usage is billed out of CABS. UNE Platform can contain both elements and will be reported separately, if applicable.

Billing — Measure 29

Title: Accuracy of Usage Feed

Area	Requirement Description
Description:	Measures the completeness of content, accuracy of information and conformance of formatting of the records the ILEC transmits to the CLEC in the reporting period. *Note: This data will be collected by CLECs and reported by the ILECs.*
Method of Calculation:	((Number of Total Correct Usage Records Processed in the Reporting Period That Reflected Complete Information Content and Proper Formatting) / (Total Number of Usage Records Received and Processed)) x 100 *Note: Total usage records includes detail data records, headers and trailers*
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate
Report By:	Total Records
Geographic Level:	Statewide
Measurable Standard:	Benchmark for Pacific Bell and GTE *Parties agree that data will be collected for this measure and the appropriate benchmark discussed at next Performance Measurement Plan Review or after three months of data are available, which ever occurs first.*
Business Rules:	• Report will be by calendar month • Usage files included in the reporting month will be those processed by the CLEC in that month • Usage feed will include Resale, UNE and Meet Point Billing usage • Results will be supplied by the CLEC to the ILEC by the 7th calendar day by 7p.m. (EST) after the end of the month under report. If no data is received by the ILEC from the CLEC by required date, no results will be reported by the ILEC for the CLEC for that reporting month. Data must be supplied by the CLEC to the ILEC in the agreed to format, at minimum including data for the numerator, denominator and the calculated result.

	• If the data received by the ILEC from the CLEC are incomplete or corrupted, the ILEC will return the data file to the CLEC. The ILEC will have 12 hours after the receipt of the monthly results from a CLEC to validate the accuracy and completeness of the file and return incomplete and/or corrupted files to the CLEC for correction. The CLEC has until the 9th calendar day at 7p.m. (EST) to re-submit the file to the ILEC for inclusion in the monthly reported results. • Usage files by the ILEC will be considered non-compliant if the ILEC has changed its file criteria without providing the CLEC notice of the change 60 days prior to implementation of changes resulting from modifications to the industry format standards or 30 days prior to implementation of changes to internal ILEC format standards. For changes to internal ILEC format standards, a CLEC may request that the implementation of the change be delayed up to 30 days to allow the CLEC a 60 day internal to implement the change in its systems. This request from the CLEC must be submitted in writing to ILEC prior to the implementation of the change. • Changes to the ILEC-specific implementation guide and the ILEC reference table shall not constitute valid criteria for the purpose of determining the accuracy of a mechanized bill unless notice of the change has been provided through an agreed-upon medium for the minimum notice period. The layout of the records exchanged between companies shall be the EMI record as described in the current edition of the EMI manual published by ATIS on behalf of the Ordering and Billing Forum, as supplemented by GTE's or Pacific Bell's specific requirements. This will include record length, field descriptions, and dataset characteristics. • Validation of accuracy and completeness of the files will be accomplished by means of pack invoice checking for proper sequencing. Further validation will occur by balancing of the record count and revenue total contained in the pack trailer to the detail records. • A record is correct if it is of the correct length, all of its fields are of correct length and mode (alpha or numeric), and it is a valid EMI record type. • A header is correct if: 1) the invoice number is correct if it is of proper sequence (the sequence is 1 greater than the previous header invoice number or it is 1 if the previous sequence was 99); 2) the trailer count and the count of detail records agree and ; 3) the trailer revenue total agrees with the total of the revenue fields within each detail record within the pack.

Notes:	• The ILEC will have the right to audit the CLECs' data collection and reporting process subject to the same notice requirements that would apply to a CLEC audit of ILEC data. • The ILEC can request the CLEC supply the raw data used to compile the monthly results subject to the same notice requirements that would apply to the ILEC's provision of raw data. • Raw data includes header, trailer and detail records, for the report period in question.

<u>Billing</u> **Measure 30**

Title: Wholesale Bill Timeliness

Area	*Requirement Description*
Description:	This measure captures the elapsed number of calendar days between the scheduled close of a Bill Cycle and the ILEC's successful transmission of the associated invoice to the CLEC.
Method of Calculation:	(Count of Invoices Transmitted by ILEC in 10 calendar days from the scheduled Bill Cycle Close*/Total Count of Invoices Transmitted in Reporting Period) X 100 *Bill Cycle Close = Bill Date
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, and by ILEC Affiliates
Report By:	• Resale • UNE (IntraLATA and InterLATAcombined) • Facilities/Interconnection
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell and GTE:** **Benchmark:** • Standard – 99% within 10 calendar days
Business Rules:	• Includes only mechanized bills. • Excludes paper bill, magnetic bill, CD ROM bill or Custom Bill diskette bill.
Notes:	• GTE legacy system billing data feeds do not support the disaggregation of UNE and Resale major service group types. GTE will report the results for Resale and UNE service group types as a total result.

Report Requirements

Billing **Measure 31**

Title: Usage Completeness

Area	*Requirement Description*
Description:	Measures the percentage of usage charges appearing on the correct bill.
Method of Calculation:	(Count of usage charges on the bill that were recorded within last 30 days / total count of usage charges on the bill) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies)and by ILEC Affiliates
Report By:	• Resale • UNE (IntraLATA and InterLATAcombined) • Facilities/Interconnection
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell and GTE:** **Parity for Resale and UNE** **Benchmark for Facilities/Interconnection** • Standard - 95%
Business Rules:	• Excludes summarized charges
Notes:	• For Pacific Bell, for CABS billed charges (UNE and Facilities/Interconnection), dataset will be defined as charges occurring in past 30 days and processed within 3 calendar days of the end of the month. • GTE legacy system billing data feeds do not support the disaggregation of UNE and Resale major service group types. GTE will report the results for Resale and UNE service group types as a total result.

Billing **Measure 32**

Title: Recurring Charge Completeness

Area	*Requirement Description*
Description:	Measures the percentage of fractional recurring charges appearing on the correct bill.
Method of Calculation:	**Pacific Bell:** (Count of fractional recurring charges that are on the correct bill* / total count of fractional recurring charges that are on the bill) x 100 *Correct bill = next available bill **GTE:** (Dollar amount of fractional recurring charges that are on the correct bill*/ total dollar amount of fractional recurring charges that are on bill) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and by ILEC Affiliates
Report By:	• Resale • UNE (IntraLATA and InterLATA combined) • Facilities/Interconnection
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell:** **Parity for Resale and UNE POTS** **Benchmark for Facilities/Interconnection and UNE Specials** • Standard – 90% **GTE:** **Parity for Resale and UNE** **Benchmark for Facilities/Interconnection** • Standard – 90%
Business Rules:	• The effective date of the recurring charge must be within one month of the bill date for the charge to appear on the correct bill. • Excludes late charges resulting from externally mandated billing changes that the ILEC can not reasonably implement in a timely manner.
Notes:	• GTE will compare CLEC results to a statistically valid sample of GTE results. • Pacific will continue to report this measure until sixty days following the implementation of Measure 35.

OSS OII Performance Measurements
Report Requirements

<u>Billing</u> Measure 33

Title: Non-Recurring Charge Completeness

Area	*Requirement Description*
Description:	Measures the percentage of non-recurring charges appearing on the correct bill.
Method of Calculation:	**Pacific Bell:** (Count of non-recurring charges that are on the correct bill* / total count of non-recurring charges that are on the bill) x 100 *Correct bill = next available bill **GTE**: (Dollar amount of non-recurring charges that are on the correct bill */ total dollar amount of non-recurring charges that are on bill) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies)and by ILEC Affiliates
Report By:	• Resale • UNE (IntraLATA and InterLATAcombined) • Facilities/Interconnection
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell:** **Parity for Resale and UNE POTS** **Benchmark for Facilities/Interconnection and UNE Specials** • Standard - 90% **GTE:** **Parity for Resale and UNE** **Benchmark for Facilities/Interconnection**: • Standard – 90%
Business Rules:	• The effective date of the non-recurring charge must be within one month of the bill date for the charge to appear on the correct bill. • Excludes late charges resulting from externally mandated billing changes that the ILEC can not reasonably implement in a timely manner.
Notes:	• Pacific will continue to report this measure until sixty days following the implementation of Measure 35.

Billing **Measure 34**

Title: Bill Accuracy

Area	*Requirement Description*
Description:	Measures the percentage of the total bill amount that is not adjusted by correcting service orders or adjustments for the month.
Method of Calculation:	(Total monies billed without corrections/total monies billed) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and by ILEC Affiliates

Report By:	• Resale
	  • Usage
	  • Recurring Charges
	  • Non-Recurring Charges
	• UNE (IntraLATA and InterLATA combined)
	  • Usage
	  • Recurring Charges
	  • Non-Recurring Charges
	• Facilities/Interconnection
	  • Usage
	  • Recurring Charges
	  • Non-Recurring Charges

Geographic Level:	Statewide

Measurable Standard:	**Pacific Bell:**
	Parity for Resale and UNE POTS
	Benchmark for Facilities/Interconnection and UNE Specials
	  • Standard - 95%
	
	GTE:
	Benchmark for Resale and UNE:
	  • Standard - 97%
	Benchmark for Facilities/Interconnection:
	  • Standard - 95%

Business Rules:	• Excludes late charges resulting from externally mandated billing changes that the ILEC can not reasonably implement in a timely manner.
Notes:	• GTE legacy system billing data feeds do not support the disaggregation of UNE and Resale major service group types. GTE will report the results for Resale and UNE service group types as a total result.

OSS OII Performance Measurements
Report Requirements

Provisioning **Measure 35**

Title: Timeliness of Billing Completion Notices - Pacific Bell Only

Area	*Requirement Description*
Description:	Measures the percent of completed orders that had a billing completion notice sent to the CLEC in 3 business days.
Method of Calculation:	Interim Method of Calculation: Sum (Number of Orders Completed in Billing Systems within 3 Business Days) / (Number of Orders Completed) x 100 As of TBD Date: Sum (Number of Billing Completion Notices Sent to CLEC within X Business Days after Work Completion) / (Number of Orders Completed) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, and by ILEC Affiliates
Reported By:	
Geographic Level:	Statewide
Measurable Standard:	**Benchmark** • Standard - 95% in 3 business days
Business Rules:	• Excludes weekends and ILEC published holidays.
Notes:	• Until the billing completion notice process has been developed Pacific will report the percentage of orders completed in the billing systems within 3 business days.

OSS OII Performance Measurements
Report Requirements

Billing **Measure 36**

Title: Accuracy of Mechanized Bill Feed

Area	*Requirement Description*
Description:	Measures the percentage of mechanized bill feeds that are accurately passed to the CLEC in the reporting period. *Note: This data will be collected by CLECs and reported by the ILECs.*
Method of Calculation:	BOS-BDT Format: (Total # of correct records + correct trailers balanced to count of records that passed / Total # of records + trailers processed in that reporting period) x 100 EDI Format: (Total # of correct segments +correct bills + correct transmissions that passed / Total # of records + bills + transmissions processed in that reporting period) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate
Report By:	BOS-BDT format and EDI format, as supplemented by GTE's or Pacific Bell's specific requirements.
Geographic Level:	Statewide
Measurable Standard:	Benchmark for Pacific Bell and GTE *Parties agree that data will be collected for this measure and the appropriate benchmark discussed at next Performance Measurement Plan Review or after three months of data are available, which ever occurs first.*

Business Rules:	
Business Rules:	• Report will be by calendar month • Transmissions included in the reporting month will be those processed by the CLEC in that month. Usage feed will include Resale, UNE and Meet Point Billing usage • Results will be supplied by the CLEC to the ILEC by the 7th calendar day by 7p.m. (EST) after the end of the month under report • If no report data is received by the ILEC from the CLEC by required date, no results will be reported by the ILEC for the CLEC for that reporting month. • Report Data must be supplied by the CLEC to the ILEC in the agreed to format, at minimum including data for the numerator, denominator and the calculated result. • If the report data received by the ILEC from the CLEC are incomplete or corrupted, the ILEC will return the data file to the CLEC. The ILEC will have 12 hours after the receipt of the monthly results from a CLEC to validate the accuracy and completeness of the file and return incomplete and/or corrupted files to the CLEC for correction. The CLEC has until the 9th calendar day at 7p.m. (EST) to re-submit the file to the ILEC for inclusion in the monthly reported results. • Mechanized bill feed transmissions by the ILEC will be considered non-compliant if the ILEC has changed its transmission criteria without providing the CLEC notice of the change 60 days prior to implementation of the change. • Changes to the ILEC-specific implementation guide and the ILEC reference table shall not constitute valid criteria for the purpose of determining the accuracy of a mechanized bill unless notice of the change has been provided through an agreed-upon medium 60 days prior to the implementation of changes resulting from modifications to the industry format standards or 30 days prior to implementation of changes to internal ILEC format standards. For changes to internal ILEC format standards, a CLEC may request that the implementation of the change be delayed up to 30 days to allow the CLEC a 60 day internal to implement the change in its systems. This request from the CLEC must be submitted in writing to ILEC prior to the implementation of the change. • A record is accurate if the billing data meets the published specifications meaning that each field of each record is of proper length and style (numeric or alpha), and it is a valid BOS-BDT or EDI file type. • A BOS-BDT record is accurate if a 99-99-99 record is included with every transmission. • A record is accurate if the bill format complies with both X12 industry guidelines and the ILEC-specific implementation guide. • A record is accurate if the codes contained I the transmission agree with the codes contained in the ILEC Reference Table • A record is accurate if the billed service type matches the service types that have been communicate tot he CLEC. • An EDI transmission is accurate if the enveloping starting segments provide accurate send/receive information and the envelope ending segments provide accurate counts.

Notes:	• BOS-BDT and EDI Billing data is considered compliant if they meet published specifications. This means that each field of each record is of proper length and style (numeric or alpha). • The ILEC will have the right to audit the CLECs' data collection and reporting process subject to the same notice requirements that would apply to a CLEC audit of ILEC data. • The ILEC can request the CLEC supply the raw data used to compile the monthly results subject to the same notice requirements that would apply to the ILEC's provision of raw data.

OSS OII Performance Measurements Report Requirements

Database Updates **Measure 37**

Title: Database Update Interval - Pacific Bell Only

Area	*Requirement Description*
Description:	Measures the average time to update databases. Reported for: • ***DA/Listings Database*** • ***LIDB (service order generated updates only)***
Method of Calculation:	**Parity Sub-measures (Service Order generated updates)** [(Completion Date & Time) – (Update Submission Date & Time)] / Count of Updates Completed in Reporting Period **Benchmark Sub-measures (Direct gateway updates)** [(Count of updates completed within 8 days)/ (Total Updates completed with in the Reporting Period)] x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate , by ILEC (if analog applies) and by ILEC Affiliates
Report By:	• Service Order generated updates • Direct gateway input
Geographic Level:	Statewide
Measurable Standard:	**Parity for service order generated updates** **Benchmark for direct gateway input updates** • Standard - 95% in 8 calendar Days
Business Rules:	
Notes:	• CLECs reserve the right to request additional databases be included in this measure. .

Database Updates — Measure 38

Title: Percent Database Accuracy - Pacific Bell Only

Area	Requirement Description
Description:	Measures the percentage of database updates completed without error. Reported for: • ***911 Databases*** • ***DA/Listings Database*** • ***LIDB***
Method of Calculation:	((Count of Updates Completed without error) / (Count of Updates Completed)) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and by ILEC Affiliates
Report By:	**DA/Listings**: • Service Order generated updates • Direct gateway input **E911 Database:** • Service Order generated updates • Direct gateway input **LIDB Database** • Service Order generated updates
Geographic Level:	Statewide
Measurable Standard:	Parity for service order generated updates Direct Gateway Input
Business Rules:	• Excludes CLEC caused errors
Notes:	• CLECs reserve the right to request additional databases be included in this measure. • Pacific Bell shall report information on direct gateway updates as a special report until Emergency 911/Listings Fix-It Team completes its work.

Database Updates **Measure 39**

Title: E911/911 MS Database Update

Area	*Requirement Description*
Description:	Measures the percentage of E911/911database updates completed within 48 hours.
Method of Calculation:	(Number of valid records updated within 48 hours / Total number of valid records updated) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate, by ILEC (if analog applies) and by ILEC Affiliates
Report By:	• Service order generated updates (Pacific Bell Only) • Direct gateway input updates
Geographic Level:	Statewide
Measurable Standard:	**Pacific Bell** Parity for service order generated updates **Pacific Bell and GTE:** Direct gateway input Standard - 48 hours
Business Rules:	• For service order generated updates, 48 hour interval begins when service order is completed in SORD (Pacific Bell) • For direct gateway updates, the processing interval is measured from the time the update enters the gateway until it posts in the 911 database. If the update rejects, the new interval starts when the update is re-submitted to the gateway.
Notes:	

OSS OII Performance Measurements Report Requirements

Collocation **Measure 40**

Title: Time to Respond to a Collocation Request

Area	*Requirement Description*
Description:	Measures the interval it takes an ILEC takes to respond to a CLEC's collocation request.
Method of Calculation:	Space Availability (# of Requests Completed in 15 Calendar Days Interval) / (Count of Requests Completed in Reporting Period) x 100 Price and Schedule Quote (# of Requests Completed in 30 Calendar Days Interval) / (Count of Requests Completed in Reporting Period) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate and by ILEC Affiliates
Report By:	• All Collocation • Space Availability • Price and Schedule Quote
Geographic Level:	Statewide
Measurable Standard:	**Space Availability -** Standard -100% in 15 calendar days **Price and Schedule Quote -** Standard - 100% in 30 calendar days

<table>
<tr><td>Business Rules:</td><td>• Excludes orders canceled by CLEC
• If the CLEC makes a change to size, location, additional AC or DC or HVAC, in their application within 15-day period or after the 15 day period, the 15-day clock is restarted from the revised application receipt date

Following are the types of changes that trigger the restarting of the 15 day clock:

• Power Upgrades - Increasing the DC power by adding a generator, rectifiers, batteries; changing power feeds; or installing a new service entrance from the electrical utility.
• HVAC Upgrades - Changing the existing cooling unit to a larger one; adding an additional cooling unit; or replacing the existing HVAC duct system to obtain additional capacity from existing units.
• Major Building Modifications - Construction activity that is required to convert space that is not suitable for housing telecommunications equipment (administrative and unconditioned space) into space that is suitable for telecommunications equipment and meets local building code. Examples of Major Building Modifications construction activities are as follows:
1. Asbestos abatement on a room or floor of a building
2. Construction of new interior partitions (walls) and doors to accommodate new HVAC system
3. Construction required to accommodate restroom access or modifications per code.
4. Construction or modification of building to facilitate proper emergency egress from the space per code.
5. Electrical wiring of space per code requirements.

• For cageless collocation, if more than 10 collocation requests are submitted per region by one CLEC within 10 calendar days, the response interval for each additional 10 requests (by region) will extend by 10 calendar days. (Pacific Bell only)</td></tr>
<tr><td>Notes:</td><td>• Interval for both sub-measures to begin upon receipt of valid request per published ILEC guidelines.
• If time intervals for new or augmented collocation installations are adopted in any future Local Competition proceeding, these time intervals shall supercede the benchmarks set under this measure and shall be measured at 100% average response time. Pacific Bell/GTE shall file by Advice Letter a compliance filing to incorporate any new requirements adopted in the Local Competition proceeding.</td></tr>
</table>

OSS OII Performance Measurements
Report Requirements

Collocation — Measure 41

Title: Time to Provide a Collocation Arrangement

Area	Requirement Description
Description:	Measures the interval it takes an ILEC to complete (build) a collocation arrangement.
Method of Calculation:	(# of Collocation Arrangements Completed in "X" Interval) / (Total Number of Collocation Arrangements Completed During the Reporting Period) x 100
Report Period:	Monthly
Report Structure:	Individual CLEC, CLECs in the aggregate and by ILEC Affiliates

Report By:	• All Collocation
	    • New
	        • Cageless
	    • Augment
	        • Cageless

Geographic Level:	Statewide
Measurable Standard:	**Benchmark for Pacific Bell:** • New - 100% compliance within time intervals set in its tariffs • Augmentation - 100% in 80 calendar days **Benchmark for GTE:** • New - 90% compliance within 90 calendar days • Augmentation - 100% in 80 calendar days

Business Rules:	• Excludes orders canceled by CLEC • Excludes CLEC requested due dates greater than the standard interval. • Applies to all requests for physical collocation space. Interval begins when ILEC approves the application and has received, from CLEC, financial payment or bond. • For cageless collocation, if more than 10 collocation arrangements are requested per region by one CLEC within 10 calendar days, the construction interval for each additional 10 requests (by region) will extend by 10 calendar days.(Pacific Bell only) • A change in a collocation request shall not trigger a restarting of the clock on the collocation interval. If, however, a CLEC delays the collocation installation, the collocation interval shall be increased by the number of days of CLEC delay (resulting in an adjusted interval). If the ILEC completes the requisite installation by the adjusted interval, it will have met its obligation under Measure 41.(Pacific Bell only).
Notes:	If time intervals for new or augmented collocation installations are adopted in any future Local Competition proceeding, these time intervals shall supercede the benchmarks set under this measure and shall be measured at 100% average response time. Pacific Bell/GTE shall file by Advice Letter compliance filing to incorporate any new requirements adopted in the Local Competition proceeding.

Interfaces **Measure 42**

Title: Percentage of Time Interface is Available

Area	*Requirement Description*
Description:	Measures percent of time OSS interface is available compared to scheduled availability.
Method of Calculation:	[(Number of Scheduled Interface Available Hours) - (Number of Unscheduled Interface Unavailable Hours)] / Scheduled System Available Hours) x 100
Report Period:	Monthly
Report Structure:	CLECs in the aggregate, by ILEC (if analog applies), ILEC Affiliate
Reported By:	By interface type for all interfaces accessed by CLECs (e.g., pre-ordering, ordering, and maintenance)
Geographic Level:	Statewide
Measurable Standard:	**Parity for Pacific Bell for interfaces used by both ILEC and CLEC** **Benchmark for Pacific Bell (for all otherinterfaces)and GTE (all interfaces)** • Standard – 99.25%
Business Rules:	• Outage hours are obtained from outage reports • Any change requests for extended availability during the reporting period are added to the scheduled hours.
Notes:	• GTE captures data on a nationwide basis and reports national results at a state level.

Interfaces **Measure 43**

Title: **MEASURE DELETED**

Area	*Requirement Description*
Description:	*Measure deleted - process is parity by design.*
Method of Calculation:	
Report Period:	
Report Structure:	
Reported By:	
Geographic Level:	
Measurable Standard:	
Business Rules:	
Notes:	

OSS OII Performance Measurements
Report Requirements

Interfaces **Measure 44**

Title: Center Responsiveness

Area	*Requirement Description*
Description:	Measures the average time it takes the ILEC's work center to answer a call.
Method of Calculation:	Sum (Date and Time of Call answer - Date and Time of Call Receipt) / (Total calls answered by center))
Report Period:	Monthly
Report Structure:	CLECs in the aggregate, and by ILEC (if analog applies)
Report By:	• ILEC Ordering Center • ILEC Repair Center • ILEC Provisioning Center (Pacific Bell)
Geographic Level:	Statewide
Measurable Standard:	**Repair Centers** Parity - Pacific Bell Benchmark – GTE Standard – average 17 seconds **Benchmark for Pacific Bell and GTE (Ordering Centers)** Standard – average 15 seconds (Pacific Bell) Standard – average 17 seconds (GTE) **Benchmark for Pacific Bell Provisioning Center** Standard - average of 90 seconds
Business Rules:	
Notes:	• Measured by individual queue, if applicable, in each ILEC center. • GTE captures data on a nationwide basis and reports national results at a state level. • GTE reports two repairs centers: 1) Designed Engineered Services; and 2) Non-designed (Non-Engineered) Services

REPORTING PROCESS

Except as otherwise provided, performance reports will be provided to the CLECs and the Public Utilities Commission by the fifteenth calendar day of the month succeeding the reporting period. The reporting period is the calendar month, unless otherwise noted. Reporting will be activity based , i. e. where there is reportable data for the CLEC.

For those measures where results appear to be statistically less than parity or not meeting the benchmark level, the ILEC will perform analysis of the data if requested by the CLEC. This analysis will detail the underlying causes contributing to the reported performance results. The ILEC will supply this analysis to the requesting CLEC within thirty days.

Authorized users will have access to monthly reports through an interactive website. Each CLEC will have access to its own data, aggregate CLEC data, ILEC data and ILEC Affiliate data. ILEC Affiliate data will be reported, at a minimum, separately for the ILEC Data subsidiary and all other ILEC Affiliates (in the aggregate). The ILECs will report performance measurements for transactions with their affiliates and make those data available to all CLECs who have filed non-disclosure documents like those filed by Pacific Bell and GTE with regard to CLEC data. The Public Utilities Commission will have access to reports for all entities, including ILEC Affiliate data. ILEC Affiliate data will not be included in CLEC aggregate data.

In addition to the performance measure results themselves, the raw data supporting the results, for the current and prior month, will be available to the CLECs and the Public Utilities Commission. Additional raw data will be available where measure results have been changed and the raw data has been affected. Raw data will be archived for a period of 24 months to provide an adequate audit trail and will be retained with sufficient detail so that CLECs can reasonably reconcile the data captured by the ILEC (for the CLEC) with its own internal data. Furthermore, data that relates to the ILEC's own performance would be retained, at a consistent level of disaggregation comparable to that reported for the CLECs.

ILEC will provide data which comprise the results and which are readily available from the systems which provide the reportable data. ILEC will provide PON information associated with Ordering and Provisioning measures. CLECs should request raw data on an as-needed basis. Pacific Bell will produce the current month's raw data within 15 days and the prior within 30 days. GTE will provide the requested data within 30 days.

Upon approval of the JPSA filed on July 18, 2000, Pacific will begin reporting performance reports to the CLECs and the Public Utilities Commission by the twentieth calendar day of the month succeeding the reporting period. Pacific expects to implement an upgrade to its reporting procedures that provides the CLECs with direct, real time access to their raw data electronically by the end of first quarter, 2001. In the event that Pacific does not implement such upgrade in the expected time frame, the CLECs may elect to have Pacific revert to reporting performance reports by the fifteenth of the month. In the interim, Pacific and CLECs will meet, on or about the tenth of each month, to discuss the feasibility of shortening Pacific's response time to CLEC requests for raw data and whether allowing Pacific to report on the twentieth of the month has reduced the number of changes necessary to the website and raw data. Pacific expects the extension in reporting time to reduce changes by as much as 25%. In the event that the extension in time does not result in a reduction in changes within 90 days, Pacific will revert to reporting performance reports by the fifteenth of the month. Until Pacific implements its upgrade, CLECs may request raw data from Pacific as early as the date Pacific reports its performance reports. Pacific will provide the requested raw data for the current reported month within fifteen days and for prior months within 30 days (or less upon agreement of the parties).

CALIFORNIA OSS OII PERFORMANCE MEASUREMENTS

SERVICE ORDER TYPES

- **New Service Installations**
- **Service Migrations without Changes**
- **Service Migrations with Changes**
- **Move and Change activities**
- **Feature Changes**
- **Service Disconnects**

AUDITING

Initial Audit:
(See prior versions of the JPSA for discussion on Initial Audit).

Annual Audits:
A comprehensive Annual Audit will be conducted of the ILECs' reporting procedures and reportable data. The Annual Audit will include all systems, processes and procedures associated with the production and reporting of performance measurement results, except as noted below A Joint Steering Committee ("Committee") comprised of ILEC and CLEC representatives will be responsible for:

1. Jointly defining the Request for Proposal;
2. Jointly selecting a third party auditor;
3. Determining the scope and timing of the Annual Audit;
4. Providing guidance to the auditor, as requested; and
5. Reviewing the auditor's compliance with the Request for Proposal.

The Committee will convene every six months to discuss the Annual Audit. In the event that the Committee cannot agree on defining the Request for Proposal, selecting an auditor, or determining the scope or timing of the Annual Audit, the parties agree to submit their disputes to the American Arbitration Association ("AAA") for expedited resolution. The AAA shall have discretion to award arbitration costs, excluding attorneys fees, to the prevailing party.

At its completion, the ILEC shall submit its annual comprehensive audit to the Commission, and distribute copies (which include only non-proprietary information) to parties on the OSS OII service list.

No Annual Audit shall commence within 12 months of the commencement of the previous Annual Audit. Notwithstanding any other provisions herein, the scope of the Annual Audit shall not exceed the previous 12 months. In addition, at least one comprehensive Annual Audit will be conducted every three years.

The costs of the Annual Audit will be divided 50% to the ILEC and 50% to the CLECs, in the proportion of each individual CLEC's volume to the aggregate CLEC volume. Volume for purposes of this allocation will be the number of local exchange lines, interconnection/interoffice trunks ('trunks"), circuits, and UNEs (as reported in the denominator of Measure 19, the "Customer Trouble Report Rate" measure) in service in the third reported month prior to the commencement of the Annual Audit. In order to assign weight to the different local exchange lines/trunks/circuits and UNEs reported in Measure 19, the Committee shall develop and approve a conversion table based on a standard unit of weight, likely using a DS-0 equivalency, including appropriate consideration for collocation; provided, the ILEC shall not in any event have an obligation to provide data or perform calculations that are not part of its normal data reporting systems.

The estimated cost of the Annual Audit (based on the chosen vendor's response to the Request for Proposal) will be paid into escrow by the ILEC and the CLECs a reasonable period of time before the commencement of the Annual Audit and shall be a prerequisite for the commencement of the Annual Audit. Any disputes regarding payments owed by the respective CLECs for the Annual

Audit shall be submitted to the American Arbitration Association (“AAA”) for expedited resolution. The AAA shall have discretion to award arbitration costs, excluding attorneys fees, to the prevailing party.

In the case of GTE, when the Annual Audit is performed at the national level for systems, processes and procedures associated with the production and reporting of performance measurement results, the Annual Audit cost in California associated with the audit of GTE’s national systems, processes and procedures shall be determine on a pro-rated basis as follows: The California portion shall be based on the volume of CLEC activity in California as compared to the total CLEC volume in all GTE states. Volume for purposes of this allocation will be the number of local exchange lines, trunks, circuits, and UNEs (as reported in Measure 19) in service in third reported month prior to the commencement of the Annual Audit. Audit costs specific to California shall be shared by GTE and the CLECs as set forth in the paragraph above.

Mini – Audits:

In addition to an annual audit, Pacific Bell, GTE and CLECs agree that the CLECs would have the right to mini-audits of individual performance measures/sub-measures during the year. When a CLEC has reason to believe the data collected for a measure is flawed or the reporting criteria for the measure is not being adhered to, it has the right to have a mini-audit performed on the specific measure/sub-measure upon written request (including e-mail), which will include the designation of a CLEC representative to engage in discussions with the ILEC about the requested mini-audit. If, 30 days after the CLEC's written request, the CLEC believes that the issue has not been resolved to its satisfaction, the CLEC will commence the mini-audit upon providing the ILEC with 5 business days advance written notice. Each CLEC is limited to auditing three single measures/sub-measures during the audit year. The Mini-audit yearwill be based on a calendar year. Mini-audits cannot be requested by a CLEC while an Annual Audit is being conducted (i.e. before completion). Mini-Audits may be requested for months including and subsequent to the month in which an Annual Audit was initiated.

Mini-Audits will include all systems, processes and procedures associated with the production and reporting of performance measurement results for the audited measure/sub-measure. Mini-Audits will include two (2) months of data, and all parties agree that raw data supporting the performance measurement results will be available monthly to CLECs as described in the Reporting Process section (Section II.c) of this agreement.

No more than three (3) Mini-Audits will be conducted simultaneously unless more than one CLEC wants the same measure/sub-measure audited at the same time, in which case, Mini-Audits of the same measure/sub-measure shall count as one Mini-Audit for the purposes of this paragraph only.

Mini-Audits will be conducted by a third party auditor, selected by the same method as the selection of the auditor for the Annual Audit. The CLEC will pay for the costs of the third party auditor conducting the Mini-Audit unless the ILEC is found to be “materially” misreporting or misrepresenting data or to have non-compliant procedures, in which case, the ILEC would pay for the costs of the third party auditor. Parties agree that the issue of whether the ILEC is “materially” at fault will be based on the parameters of failure to perform: “materially” at fault means that a reported successful measure changes as a consequence of the audit to a missed measure, or there is a change from an ordinary missed measure to another category, if such exists. Each party to the

Mini-Audit shall bear its own internal costs, regardless of which party ultimately bears the costs of the third party auditor.

If, during a Mini-Audit, it is found that for more than 50% of the measures in a major service category the ILEC is "materially" at fault (i.e., a reported successful measure changes as a consequence of the audit to a missed measure, or there is a change from an ordinary missed measure to another category, if such exists), the entire service category will be re-audited at the expense of the ILEC. The major service categories for this purpose are:

- Pre-Ordering
- Ordering
- Provisioning
- Maintenance
- Network Performance
- Billing
- Database Updates
- Collocation
- Interfaces

Each Mini-Audit shall be submitted to the CLEC involved and to the Commission as a proprietary document subject to the applicable protection afforded by Commission General Order No. 66 C and California Public Utilities Code Section 583.

The ILEC will provide notification to the CLECs of any Mini-Audit requested when the request for the audit is made.

REVIEW PROCEDURES

As experience is acquired under this Partial Settlement Agreement with the new performance measurements and underlying business processes, the Parties expect to learn which measurements set forth in Section II may not have been properly defined or are more or less useful than others. The Parties also expect that experience will show whether new measurements are needed or whether certain existing measurements are not needed or require modification. Accordingly, the Parties agree to reconvene on or aroundMarch 1, 2001 to review the effectiveness of and modifications to the performance measurements approved by the Commission in this proceeding. The parties will conclude the review within 90 days of its commencement and will submit the revisions to the Partial Settlement Agreement to the Commission within the 90 day review period. In the event the Parties cannot agree on any addition, deletion or modification, they will jointly submit such dispute for resolution by the CPUC.

If, prior to the agreed-upon review date, there is consensus that one or more measures are not effective, the parties will schedule meetings to discuss modifying the measure(s) or process(es). If there is no consensus, any individual party seeking formal review by the CPUC shall give notice to the other parties of its intent to do so. The party will also describe the action it intends to take and the reason(s) for its proposed actions.

Implementation Timeline for Pacific Bell Changes to JPSA

***Note: Implementation interval begins when revised JPSA is ordered by the Commission**

Item No.	*Measure*	*Sub-Measure*	*Change*	*Date of Change**
1	1	Electronic Pre-order Queries	Measure as total transaction time	Completed
2		Electronic loop qual sub-measure	New sub-measure	Completed
3		Manual loop qualification	New sub-measure	Completed
4		CSR sub-measures	Change project limit to 50 TNs	30 Days
5	2	Projects	New sub-measure	30 Days
6		Sub-measures associated with xDSL and Line /Sharing, ISDN, channelized DS1, DS3 and Unbundled Ded. Transport (DS3)	Exclude pre-qual time	Completed
7		Held and Denied Interconnection Trunk reports	Measure at parity with retail	90 Days
8	3	Line Sharing	New sub-measure	Completed
9		Standalone Directory Listings	New sub-measure	90 Days
10		Projects	New sub-measure	30 Days
11		Sub-measures associated with xDSL and Line /Sharing, ISDN, channelized DS1, DS3 and Unbundled Ded. Transport (DS3)	Exclude pre-qual time	Completed
12	4			
13	5	"Electronic interface" disaggregation	Eliminate disaggregation	60 Days
14		"Lack of facilities and all other" disaggregation	Eliminate disaggregation	60 Days
15		2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
16	5	Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
17		UNE Platform sub-measures	New Sub-measures	90 Days
18		UNE port sub-measures	Consolidate to UNE Port (non special) and UNE Port (special)	90 Days
19		UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
20		Raw Data	Include jeopardy codes	60 Days
21	6	"Electronic interface" disaggregation	Eliminate disaggregation	60 Days
22		"Lack of facilities and all other" disaggregation	Eliminate disaggregation	60 Days
23		2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
24		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
25		UNE Platform sub-measures	New Sub-measures	60 Days
26		UNE port sub-measures	Consolidate to UNE Port (non special) and UNE Port (special)	90 Days
27		UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	60 Days
28		Raw Data	Include jeopardy codes	60 Days
29	7	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
30		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
31		UNE Platform sub-measures	New Sub-measures	90 Days
32		All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 days
33	7	UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days

34		UNE port sub-measures	Consolidate to UNE Port (non special) and UNE Port (special)	90 Days
35	8	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	90 Days
36		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
37		UNE Platform sub-measures	New Sub-measures	90 Days
38		All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 days
39		UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
40		UNE port sub-measures	Consolidate to UNE Port (special)	90 Days
41	9	Total measure	Base measures on total cutovers scheduled, not total coordinated conversion orders	Completed
42	9A	Total measure	Implement this new measure	180 Days
43	10	Total measure	Change to benchmark	Completed
44		Total measure	Exclude large ports (greater than 500 TNs)	30 Days
45	11	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
46		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
47		UNE Platform sub-measures	New Sub-measures	90 Days
48		All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 Days
49	11	UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
50		UNE port sub-measures	Consolidate to UNE Port (non special) and UNE Port (special)	90 Days
51	12	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
52		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
53		UNE Platform sub-measures	New Sub-measures	90 Days
54		All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 Days
55		UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
56	13	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
57		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
58		UNE Platform sub-measures	New Sub-measures	90 Days
59		All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 Days
60		UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
61	14	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
62		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
63		UNE Platform sub-measures	New Sub-measures	90 Days
64		All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 Days
65	14	UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
66		UNE port sub-measures	Consolidate to UNE Port (non special) and UNE Port (special)	90 Days

67	15	UNE Loop sub-measure	Include central office wiring code troubles in retail analog	Completed
68	15A	Total measure	Implement new measure	60 Days
69	16	UNE Loop sub-measure	Include central office wiring code troubles in retail analog	Completed
70		Total measure	Redefine measure to only include special service orders	30 Days
71	17	Total measure	Implement measure to only include non-special service orders	30 Days
72	18	Fully electronic sub-measures	Eliminate fallout results from sub-measures	30 Days
73		Fully electronic fallout sub-measures	Implement new sub-measures	30 Days
74	35	Total measure	Implement new measure (Phase 1) Implement billing notification process (Phase 2)	90 Days TBD
75	19, 20, 21, 23	2/4w (5.5db) analog loop	Eliminate disaggregation -combine with basic (8db) UNE loops	60 Days
76		Advanced Services sub-measures (UNE Subloop, Dark Fiber, EELs)	New sub-measures	90 Days
77		UNE Platform sub-measures	New Sub-measures	90 Days
78		All UNE Loop sub-measures	Exclude feature only orders from Retail analog	60 Days
79		UNE Ded. Transport sub-measure	Disaggregate by DS1 and DS3	30 Days
80	19, 20, 21, 23	UNE port sub-measures	Consolidate to UNE Port (non special) and UNE Port (special)	90 Days
81		UNE Loop sub-measure	Include central office wiring code troubles in retail analog	Completed
82	22	All UNE Loop submeasures	Exclude feature only orders from Retail analog	60 Days
83		UNE Loop sub-measure	Include central office wiring code troubles in retail analog	Completed
84	24	Total measure	Report at statewide level and make available detail at trunk group level for not meeting 2% or less blocking level	Completed
85	25	Total measure	Report at statewide level and make available detail at trunk group level for not meeting parity	Completed
86		Total measure	Exclude performance failures caused by CLEC not completing growth provisioning on time	30 Days
87	26	Total Measure	Exclude performance failures where no test number provided or interconnection facilities not installed	30 Days
88	27	Total Measure	Eliminate measure	30 Days
89	28	Jointly provided switched access sub-measure	Change from benchmark to parity comparison	30 Days
90	29, 36	Total measure	Report results using new business rules	Completed CLEC Provided Data
91	31	UNE and Facilities/Interconnect sub-measures	Redefine data collection period to collect all usage data occurring in past 30 days and processed within 3 business days of the end of the month	180 Days

92	32,33	Total measure	Exclude late charges resulting from mandated billing changes that cannot be implemented in a timely manner	30 Days
93	34	Total measure	Exclude late charges resulting from mandated billing changes that cannot be implemented in a timely manner	30 Days
94	37, 38	LIDB sub-measure (service order generated updates)	Implement new sub-measure	180 Days
95	43	Total Measure	Eliminate measure	Completed
96	44	ILEC Prov. Center sub-measure	Implement new sub-measure	Completed

Implementation Timeline for GTE Changes Due To JPSA Changes

Item No.	*Measure*	*Sub-Measure (From 9-7-99 JPSA)*	*Change*	*Date of Change*[7]
1	1	Average Response Time OSS	New Rule: "Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours"	Complete
2		Average Response Time- Legacy (GTE and CLEC)	New Rule: "Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours"	Complete
3		Average Response Time- CSR	New Rule: "Clock Hours Excludes Non-Business Days"	120 Days
4		Average Response Time- CSR	New Rule: "Elapsed Time For Manual Processes Tracked During Published Business Hours"	Complete
5		Average Response Time- CSR WISE	New Rule: "Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours"	Complete
6		Average Response Time- CSR Fully Electronic	New Rule: "Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours"	Complete
7		Loop Qualification Transaction Time	New Rule: "Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours"	Complete
8		Average Response Time OSS	Change "Number of Queries Submitted" to "Number of Queries Returned"	30 Days
9		Average Response Time- Legacy (GTE and CLEC)	Change "Number of Queries Submitted" to "Number of Queries Returned"	30 Days
10		Average Response Time- CSR	Replace "X Business" with "24 Clock"	120 Days
11		Average Response Time- CSR	Change "Number of Queries Submitted" to "Number of Queries Returned"	30 Days
12		Average Response Time- CSR WISE	Replace "X Business" with "3 System"	120 Days
13		Loop Qualification Transaction Time	Sum ((Query Response Date and Time) - (Query Submission Date and Time)) / (Number of Queries Returned in Reporting Period)	30 Days
14		Average Response Time- Legacy (GTE and CLEC)	Insert "To Legacy System" In Denominator	30 Days
15		Average Response Time OSS	Legacy Result + 5 Seconds	150 Days
16		Average Response Time- CSR	Change to "98% in 24 Hours"	120 Days
17		Average Response Time- CSR WISE	Change to "98% in 3 System Hours"	120 Days
18		Average Response Time OSS	Title should be Pre-Order Query Transaction Time	30 Days
19		Average Response Time- Legacy (GTE and CLEC)	Title should be Legacy System Transaction Time	30 Days
20		Average Response Time- CSR	Replace Title with "Response Time- Manual CSRs"	30 Days
21		Average Response Time- Legacy (GTE and CLEC)	Display Legacy Results Only In GTE Columns (No Information To Be Displayed Under CLEC-Related Columns)	30 Days
22	2	Average FOC Notice Interval	1) Excludes delays caused for customer reasons; 2) Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours; 3) Business day = Monday through Friday, excluding weekends and ILEC published holidays.	150 Days

[7] **"Date of Change" field explanation. Assuming a PUC order on 7/31/2000, 30 Days=Aug. report month, 60 Days = Sept. report month, 90 Days = Oct. report month, 120 Days = Nov. report month, 150 Days = Dec. report month.**

Item No.	*Measure*	*Sub-Measure (From 9-7-99 JPSA)*	*Change*	*Date of Change*[7]
23		Average LSC Notice Interval	1) Excludes delays caused for customer reasons; 2) Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours; 3) Business day = Monday through Friday, excluding weekends and ILEC published holidays.	150 Days
24		Average FOC Notice Interval	Change benchmark for Interconnection Trunks from "Average 5 Days" to "Average 5 Business Days"	150 Days
25		Average LSC Notice Interval	Standalone Directory Listings as a separate disaggregation.	120 Days
26	3	Average Reject Notice Interval	New Rules: 1) "Elapsed Time For Fully Electronic Sub-Measures Tracked During Published System Hours;" 2) Business day = Monday through Friday, excluding weekends and ILEC published holidays; 3) Excludes delays caused for customer reasons.	150 Days
27		Average Reject Notice Interval	Clarify "Mechanized" denominator calculation from "# of Orders Rejected" to "(Number of Mechanized Orders Rejected in the Reporting Period)"	30 Days
28		Average Reject Notice Interval	Clarify "Manual" denominator from "Number of Faxes Submitted" to "Number of Faxes Rejected"	30 Days
29		Average Reject Notice Interval	Add UNE line sharing (total of conditioned and non-conditioned) and stand alone directory listings.	120 Days
30	4	Percentage of Flow Through Orders Currently Programmed	Add "Excludes orders rejected due to CLEC caused syntax errors, but does not exclude CLEC caused content errors."	150 Days
31		Percentage of Flow Through Orders	Add "Excludes orders rejected due to CLEC caused syntax errors, but does not exclude CLEC caused content errors."	150 Days
32		Percentage of Flow Through Orders Currently Programmed	Change numerator from "mechanized orders" to "electronically received orders" and change denominator from "mechanized service request" to "electronically received orders."	30 Days
33		Percentage of Flow Through Orders	Change numerator from "mechanized orders" to "electronically received orders" and change denominator from "mechanized service request" to "electronically received orders."	30 Days
34		Percentage of Flow Through Orders	Remove SGT/SOT requirements; replace with "All electronically received orders."	120 Days
35		Percentage of Flow Through Orders Currently Programmed	Remove SGT/SOT requirements; replace with "All electronically received orders programmed to flow through."	120 Days
36	5	Percentage of Orders Jeopardized	Raw data will include jeopardy codes- LSRs.	Complete
37		Percentage of Orders Jeopardized	Raw data will include jeopardy codes- ASRs.	Complete
38		Percentage of Orders Jeopardized	Remove "By electronic interface" and "By lack of facilities and all other"- LSRs.	120 Days
39		Percentage of Orders Jeopardized	Remove "By electronic interface" and "By lack of facilities and all other"- ASRs.	120 Days
40		Percentage of Orders Jeopardized	Reference SGT Table- LSRs.	150 Days
41		Percentage of Orders Jeopardized	Reference SGT Table- ASRs.	150 Days
42		Percentage of Orders Jeopardized	Change title from "Percentage of Orders (LSRs) Given Jeopardy" to "Percent of Orders Jeopardized"- LSRs.	30 Days
43		Percentage of Orders Jeopardized	Change title from "Percentage of Orders (ASRs) Given Jeopardy" to "Percent of Orders Jeopardized"- ASRs.	30 Days
44	6	Average Jeopardy Notice Interval	Raw data will include jeopardy codes.	30 Days
45		Average Jeopardy Notice Interval	Change denominator from "Order Jeopardized" to "Assignment Jeopardy Notices" for the assignment calculation.	30 Days

Item No.	Measure	Sub-Measure (From 9-7-99 JPSA)	Change	Date of Change[7]
46		Average Jeopardy Notice Interval	Remove "By electronic interface" and "By lack of facilities and all other." Note: this applies to all three "Methods of Calculation."	150 Days
47		Average Jeopardy Notice Interval	Reference SGT Table; note: SGT applies to all three "Methods of Calculation"	150 Days
48	7	Average Completed Interval	Add 1) GTE will not exclude projects; 2) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review	Complete
49		Average Completed Interval	Reference SGT Table	150 Days
50	8	Percent Completed within Standard Interval	Add 1) GTE will not exclude projects; 2) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	Complete
51		Percent Completed within Standard Interval	Remove Excludes services with flexible due date i.e., B1/R1 Service (GTE).	Complete
52		Percent Completed within Standard Interval	Reference SGT Table	Complete
53	10	PNP Network Provisioning	Change all references from PNP to LNP.	120 Days
54		PNP Network Provisioning	New business rule reads: "Provisioning failure data will be collected as follows: · Will be tracked for individual network database failures - failures to provision between the ILEC LSMS and LNP network databases (STP or SCP)."	120 Days
55		PNP Network Provisioning	Change from parity to benchmark of 2% failure.	120 Days
56	11	Percent of Due Dates Missed	Add business rules: 1) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review; 2) Excludes records only ILEC official orders.	Complete
57		Percent of Due Dates Missed	Change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Missed Appointment reason codes as diagnostic data" to "ILECs will provide disaggregation by Missed Appointment reason codes as diagnostic data upon raw data request."	30 Days
58		Percent of Due Dates Missed	Reference SGT Table	150 Days
59	12	Percent of Due Dates Missed Due to Lack of Facilities	Reference SGT Table	150 Days
60	13	Delay Order Interval to Completion Date (For Lack of Facilities)	Reference SGT Table	150 Days
61	14	Held Order Interval	Change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Missed Appointment reason codes as diagnostic data" to "ILECs will provide disaggregation by Jeopardy Code as diagnostic data upon raw data request."	30 Days
62		Held Order Interval	Reference SGT Table	150 Days
63	15	Provisioning Trouble Reports	New Business rule: Excludes new service installations. Change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request."	Complete
64	15A	Average Time To Restore Provisioning Troubles (Prior To Service Order Completion)	New Measure. Same business rules (with modifications) on PM 15 apply to PM15A.	120 Days
65		Average Time To Restore Provisioning Troubles (Prior	New Measure (Total duration of provisioning trouble measured from the time the trouble was initiated or called in to the ILEC until cleared. and verified with the CLEC)/ (Total Number of Provisioning Trouble Reports)	120 Days

Item No.	Measure	Sub-Measure (From 9-7-99 JPSA)	Change	Date of Change[7]
		To Service Order Completion)		
66		Average Time To Restore Provisioning Troubles (Prior To Service Order Completion)	New Measure Reference SGT Table; also by "Affecting Service" and Out of Service."	120 Days
67	16	Percentage Troubles in 30 days for New Orders	Change from 1) "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request;" 2) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	Complete
68		Percentage Troubles in 30 days for New Orders	Reference SGT Table	150 Days
69		Percentage Troubles in 30 days for New Orders	Change title from "New Orders" to "Designed Service Orders"	30 Days
70	17	Percentage Troubles in 7 Days for New Orders- GTE Only	Change from 1) "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request;" 2) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	Complete
71		Percentage Troubles in 7 Days for New Orders- GTE Only	Change denominator from "Total new, move and change orders" to "Total new, move and change completed orders"	30 Days
72		Percentage Troubles in 7 Days for New Orders- GTE Only	Reference SGT Table	150 Days
73	18	Average Completion Notice Interval	New rules: Completion Notices on disconnect orders are only on CLEC disconnect orders (not on ILEC retail disconnect orders) For All Other Interfaces.	Complete
74		Average Completion Notice Interval	New rules: 1) System hours will be used for fully electronic sub-measures; 2) Completion Notices on disconnect orders are only on CLEC disconnect orders (not on ILEC retail disconnect orders) for Fully Electronic.	Complete
75		Average Completion Notice Interval	Change from "Sum (# of Completion Notices Returned within “X” Interval) / (# of Orders Completed) x 100 to "(Number of Completion Notices Returned within “X” Interval) / (Number of Orders Returned Using All Other Processes) x 100 For All Other Interfaces	30 Days
76		Average Completion Notice Interval	Change from "Sum ((Date and Time of Completion Notification to CLEC) - (Date and Time of Work Completion)) / (Number of Orders Completed) to (Number of Completion Notices Returned within “X” Interval) / (Number of Orders Completed where the Completion Notice is Returned Using Electronic Process) x 100 for Fully Electronic	120 Days
77		Average Completion Notice Interval	Change from "Average Completion Notice Interval" to "Completion Notice Interval" for All Other Interfaces.	30 Days
78		Average Completion Notice Interval	Change from "Average Completion Notice Interval" to "Completion Notice Interval" for Fully Electronic.	120 Days
79		Average Completion Notice Interval	Change from "Average Completion Notice Interval (LSC)" to "Completion Notice Interval" for the WISE Web Display.	120 Days
80	19	Customer Trouble Report Rate	New business rules: 1) Excludes provisioning trouble reports; 2) Include Test okay (TOK) and Found Okay (FOK) reports; 3) change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request;" 4) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	Complete

Item No.	*Measure*	*Sub-Measure (From 9-7-99 JPSA)*	*Change*	*Date of Change*[7]
81		Customer Trouble Report Rate	Reference SGT Table	150 Days
82	20	Percentage of Customer Trouble not Resolved within Estimated Time	New business rules: 1) Include Test okay (TOK) and Found Okay (FOK) reports; 2) change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request;" 3) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review; 4) Excludes provisioning trouble reports.	Complete
83		Percentage of Customer Trouble not Resolved within Estimated Time	Reference SGT Table	150 Days
84	21	Average Time to Restore	New business rules: 1) Excludes provisioning trouble reports; 2) Include Test okay (TOK) and Found Okay (FOK) reports; 3) change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request;" 4) Results for Dark Fiber will be tracked diagnostically, until next periodic Performance Measures review.	Complete
85		Average Time to Restore	Reference SGT Table	150 Days
86	22	POTS Out of Service less than 24 Hours	Business rule change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request."	Complete
87		POTS Out of Service less than 24 Hours	Reference SGT Table	150 Days
88	23	Frequency of Repeat Troubles in 30 day period	Business rule change from "When results are less than parity for a reporting period, ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data" to "ILECs will provide disaggregation by Maintenance Disposition codes as diagnostic data upon raw data request."	Complete
89		Frequency of Repeat Troubles in 30 day period	Reference SGT Table	150 Days
90	24	Percent Blocking on Common Trunks	ILEC will make available detailed information (trunk group identifier, CLLI A, CLLI Z, blocking level) for all trunk groups not meeting 2% blocking level with the monthly report.	120 Days
91	24	Percent Blocking on Common Trunks	Remove "Includes Histogram Distribution Chart" and performance measure 24b.	120 Days
92		Percent Blocking on Common Trunks	Report by Total Trunk Groups.	120 Days
93	25	Percent Blocking on Interconnection Trunks	Add new business rule "Excludes blocking failures caused by the CLEC not completing growth trunk provisioning by scheduled due date."	120 Days
94		Percent Blocking on Interconnection Trunks	Remove: 1) Includes histogram distribution chart and move to Business Rules "2) Applies to those trunks where the ILEC has augmentation control; 3) Does not apply when trunks are provisioned as two-way trunks."	Complete
95		Percent Blocking on Interconnection Trunks	Remove "Includes Histogram Distribution Chart" and performance measure 25b.	120 Days
96		Percent Blocking on Interconnection Trunks	Report by Total trunk groups, ILEC end office to CLEC end office, and ILEC tandem to CLEC end office.	120 Days
97	26	NXX Loaded by LERG Effective Date	Add new business rule: Excludes any NXX code that cannot be completely tested because the CLEC has not provided an accurate test number or because CLEC facilities have not been installed.	Complete
98		NXX Loaded by LERG Effective Date	Add business rule: NXX activity includes additions and deletions (being returned to industry for reuse).	Complete
99	27	Network Outage	Delete PM.	30 Days

Item No.	Measure	Sub-Measure (From 9-7-99 JPSA)	Change	Date of Change[7]
		Notification		
100	30	Wholesale Bill Timeliness	Clarify with following: GTE legacy system billing data feeds do not support the disaggregation of UNE and Resale major service group types. GTE will report the results for Resale and UNE service group types as a total result.	Complete
101		Wholesale Bill Timeliness	Change “X” to "10 calendar."	30 Days
102		Wholesale Bill Timeliness	Clarify benchmark to 99% within 10 calendar days.	Complete
103	31	Usage Completeness	Clarify with following: GTE legacy system billing data feeds do not support the disaggregation of UNE and Resale major service group types. GTE will report the results for Resale and UNE service group types as a total result.	Complete
104	32	Recurring Charge Completeness	Change from "The effective date of the recurring charge must be within 30 days of the bill date for the charge to appear on the correct bill" to "The effective date of the recurring charge must be within one month of the bill date for the charge to appear on the correct bill." New business rule: "Excludes late charges resulting from mandated billing changes that the ILEC can not reasonably implement in a timely manner."	120 Days
105		Recurring Charge Completeness	Clarify calculation to "(Dollar amount of fractional recurring charges that are on the correct bill */ total dollar amount of fractional recurring charges that are on bill) x 100"	30 Days
106	33	Non-Recurring Charge Completeness	Change from "The effective date of the recurring charge must be within 30 days of the bill date for the charge to appear on the correct bill" to "The effective date of the recurring charge must be within one month of the bill date for the charge to appear on the correct bill." New business rule: "Excludes late charges resulting from mandated billing changes that the ILEC can not reasonably implement in a timely manner."	120 Days
107		Non-Recurring Charge Completeness	Clarify calculation to "(Dollar amount of non-recurring charges that are on the correct bill */ total dollar amount of non-recurring charges that are on bill) x 100"	120 Days
108	34	Bill Accuracy	Clarify with following: GTE legacy system billing data feeds do not support the disaggregation of UNE and Resale major service group types. GTE will report the results for Resale and UNE service group types as a total result; new business rule: "Excludes late charges resulting from mandated billing changes that the ILEC can not reasonably implement in a timely manner."	Complete
109	40	Time to Respond to a Collocation Request - Space Availability	If CLEC makes a change to size, location, additional AC or DC or HVAC, in their application within 15-day period, 15-day clock is restarted from revised application receipt date- Open Issue.	30 Days
110		Time to Respond to a Collocation Request - Price and Schedule Quote	Change from (# of Requests Returned in “X” Interval) / (Count of Requests Submitted in Reporting Period) x 100 to (# of Requests Completed in 30 Calendar Days Interval) / (Count of Requests Completed in Reporting Period) x 100	30 Days
111		Time to Respond to a Collocation Request - Space Availability	Change from (# of Requests Returned in “X” Interval) / (Count of Requests Submitted in Reporting Period) x 100 to (# of Requests Completed in 15 Calendar Days Interval) / (Count of Requests Completed in Reporting Period) x 100	30 Days
112		Time to Respond to a Collocation Request - Price and Schedule Quote	Clarify benchmark to 100% in 30 calendar days.	Complete
113		Time to Respond to a Collocation Request - Space Availability	Clarify benchmark to 100% in 15 calendar days.	Complete
114		Time to Respond to a Collocation Request - Price and Schedule Quote	Change title to "Time To Respond To A Collocation Request - Price and Schedule Quote"	30 Days
115		Time to Respond to a Collocation Request – Space	Change title to "Time To Respond To A Collocation Request - Space Availability"	30 Days

Item No.	*Measure*	*Sub-Measure (From 9-7-99 JPSA)*	*Change*	*Date of Change*[7]
		Availability		
116	41	Time to Provide a Collocation Arrangement – New	New business rule: Excludes CLEC requested due dates greater than the standard interval.	120 Days
117		Time to Provide a Collocation Arrangement - Augment	New business rule: Excludes CLEC requested due dates greater than the standard interval.	120 Days
118		Time to Provide a Collocation Arrangement - New	Clarify benchmark to 90% compliance within 90 calendar days.	Complete
119		Time to Provide a Collocation Arrangement - Augment	Clarify benchmark to 100% in 80 calendar days.	Complete
120		Time to Provide a Collocation Arrangement - New	Change to "Time To Provide A Collocation Arrangement - New"	30 Days
121		Time to Provide a Collocation Arrangement - Augment	Change to "Time to Provide a Collocation Arrangement - Augment"	30 Days
122	42	Percent of Time Interface is Available	Clarification: Change from ((Number of Scheduled System Available Hours) - (Number of Unscheduled System Unavailable Hours)) / Scheduled System Available Hours) x 100 to [(Number of Scheduled Interface Available Hours) - (Number of Unscheduled Interface Unavailable Hours)] / (Scheduled System Available Hours) x 100	30 Days
123		Percent of Time Interface is Available	Clarify: GTE captures data on a nationwide basis and reports national results at a state level.	Complete
124		Percent of Time Interface is Available	Clarify: change from GTE (all systems) Standard – 99.25% to GTE (All Interfaces) Standard - 99.25%	Complete
125		Percent of Time Interface is Available	Add ILEC affiliate.	Complete
126	43	Notification of Interface Outages	Delete PM.	30 Days
127	44	Center Responsiveness	Clarify GTE captures data on a nationwide basis and reports national results at a state level.	Complete
128		Center Responsiveness	Change benchmark from Standard – average 20 seconds to Standard – average 17 seconds for both repair and ordering centers.	30 Days

DEFINITION OF TERMS

TERM	DEFINITION
Automatic Location Information (ALI)	The feature of E911 that displays at the Public Safety Answering Point (PSAP) the street address of the calling telephone number. This feature requires a data storage and retrieval system for translating telephone numbers to the associated address. ALI may include Emergency Service Number (ESN), street address, room or floor, and names of the enforcement, fire and medical agencies with jurisdictional responsibility for the address. The Management System (E911) database is used to update the Automatic E911 Location Information databases.
Cageless Collocation	Shall have meaning set forth in FCC 1st Report and Order on Deployment of Wireline Services Offering Advanced Telecommunications Capability or any future, assoc. orders
Call Blocking	A condition on a telecommunications network where, due to a maintenance problem or an over capacity situation in a part of the network, some or all originating or terminating calls cannot reach their final destinations. Depending on the condition and the part of the network affected, the network may make subsequent attempts to complete the call or the call may be completely blocked. If the call is completely blocked, the calling party will have to re-initiate the call attempt.
Code Opening	Process by which new NPA/NXXs (area code/prefix) are defined, through software translations to network databases and switches, in telephone networks. Code openings allow for new groups of telephone numbers (usually in blocks of 10,000) to be made available for assignment to an ILEC's or CLEC's customers, and for calls to those numbers to be passed between carriers.
Common Channel Signaling System 7 (CCSS7)	A network architecture used to for the exchange of signaling information between telecommunications nodes and networks on an out-of-band basis. Information exchanged provides for call set-up and supports services and features such as CLASS and database query and response.
Common Transport	Trunk groups between tandem and end office switches that are shared by more than one carrier, often including the traffic of both the ILEC and several CLECs.
Completion	The time in the order process when the service has been provisioned and service.
Completion Notice	A notice the ILEC provides to the CLEC to inform the CLEC that the requested service order activity is complete.
Coordinated Customer Conversion	Orders that have a due date negotiated between the ILEC, the CLEC, and the customer so that work activities can be performed on a coordinated basis under the direction of the receiving carrier.
Customer Requested Due Date	A specific due date requested by the customer which is either shorter or longer than the standard interval or the interval offered by the ILEC.
Customer Trouble Reports	A report that the carrier providing the underlying service opens when notified that a customer has a problem with their service. Once resolved, the disposition of the trouble is changed to closed.
TERM	**DEFINITION**

Dedicated Transport	A network facility reserved to the exclusive use of a single customer, carrier or pair of carriers used to exchange switched or special, local exchange, or exchange access traffic.
Delayed Order	An order which has been completed after the scheduled due date and/or time
Directory Assistance Database	A database that contains subscriber records used to provide live or automated operator-assisted directory assistance. Including 411, 555-1212, NPA-555-1212.
Directory Listings	Subscriber information used for DA and/or telephone directory publishing, including name and telephone number, and optionally, the customer's address.
DS-0	Digital Service Level 0. Service provided at a digital signal speed commonly at 64 kbps, but occasionally at 56 kbps.
DS-1	Digital Service Level 1. Service provided at a digital signal speed of 1.544 Mbps.
DS-3	Digital Service Level 3. Service provided at a digital signal speed of 44.736 Mbps.
Due Date	The date provided on the FOC the ILEC sends the CLEC identifying the planned completion date for the order.
End Office Switch	A switch from which an end users' exchange services are directly connected and offered.
Firm Order Confirmation **(**FOC)	Notice the ILEC sends to the CLEC to notify the CLEC that it has received the CLECs service order, created a service request, and assigned it a due date.
Flow-Through	The term used to describe whether a LSR electronically is passed from the OSS interface system to the ILEC legacy system to automatically create a service order. LSRs that do not flow through require manual intervention for the service order to be created in the ILEC legacy system.
Held Order	An order for which the ILEC has issued a FOC, but whose due date has passed without it being completed.
High Bandwidth Line Sharing UNE	The frequency range above the voiceband on a copper loop facility that is being used to carry analog circuit switched voiceband transmissions.
Installation	The activity performed to activate a service.
Installation Troubles	A trouble, which is identified after service order activity and installation, has completed on a customer's line. It is likely attributable to the service activity (within a defined time period).
Inside Wiring	The telecommunications wiring located at a customer's premises that extends beyond the demarcation point.
Interconnection Trunks	A network facility that is used to interconnect two switches generally of different local exchange carriers
Interface Outage	A planned or unplanned failure resulting the unavailability or access degradation of a system.
Jeopardy	A failure in the service provisioning process which results potentially in the inability of a carrier to meet the committed due date on a service order.
Jeopardy Notice	The actual notice that the ILEC sends to the CLEC when a jeopardy condition has been identified.

DEFINITION OF TERMS

TERM	DEFINITION
Lack of Facilities	A shortage of cable facilities identified after a due date has been committed to a customer, including the CLEC. The facilities shortage may be identified during the inventory assignment process, or during the service installation process. If no facilities are available, the ILEC will issue a jeopardy.
Local Exchange Routing Guide (LERG)	A Bellcore master file that is used by the telecom industry to identify NPA-NXX routing and homing information, as well as network element and equipment designations. The file also includes scheduled network changes associated with activity within the North American Numbering Plan (NANP).
Local Exchange Traffic	Traffic originated on the network of a LEC in a local calling area that terminates to another LEC in a local calling area.
Local Number Portability	A network technology which allows end user customers to retain their telephone number when moving their service between local service providers. This technology does not employ remote call forwarding, but actually allows the customer's telephone number to be moved and redefined in the network of the new service provider. The activity to move the telephone number is called "porting."
Local Service Confirmation	OBF term for a FOC
Mechanized Bill	A bill that is delivered via electronic transmission.
Meet Point Billing	A billing arrangement used when two or more LECs jointly provide access to and from an interexchange carrier (IEC) for inter LATA traffic. This arrangement can be Single Bill, where one LEC bills the IEC on behalf of both LECs and remits payment to the other LEC or Multiple Bill, where each LEC bills their portion directly to the IEC.
Missed Commitment Notification	A notice from ILEC to inform CLEC that the committed due date on an order has been missed.
Non-Recurring Charge	A rate charged for a product or a service that is assessed on a one time basis.
NXX, NXX Code or Central Office Code	The three digit switch entity indicator that is defined by the "D", "E", and "F" digits of a 10-digit telephone number within the NANP. Each NXX Code contains 10,000 station numbers.
Permanent Number Portability (also known as Local or Long Term Number Portability)	A network technology which allows end user customers to retain their telephone number when moving their service between local service providers. This technology does not employ remote call forwarding, but actually allows the customer's telephone number to be moved and redefined in the network of the new service provider. The activity to move the telephone number is called "porting".
Physical Collocation	Shall have the meaning set forth in 47 C.F.R. Section 51.5.
Plain Old Telephone Service (POTS)	Refers to basic 2 wire analog residential and business services. Can include feature capabilities (e.g., CLASS features).

DEFINITION OF TERMS

TERM	DEFINITION
Projects	Service requests that exceed the line size and/or level of complexity which would allow for the use of standard ordering and provisioning processes. Generally, due dates for projects are negotiated, coordination of service installations/changes is required and automated provisioning may not be practical.
Provisioning Troubles	A trouble report that is opened for a customer's existing or new service for a trouble identified between the time of the service order creation to the time of order completion. Provisioning troubles that are associated with a CLECs customers include troubles that occur and are reported during the conversion of an ILEC customer to a CLEC.
Query Types	Pre-ordering information that is available to a CLEC that is categorized according to standards issued by OBF, the FCC and/or the CPUC.
Recurring Charge	A rate charged for a product or service that is assessed each successive billing period.
Reject	A status that can occur to a CLEC submitted local service request (LSR) when it does not meet certain criteria. There are two types of rejects:, syntax, which occur if required fields are not included in the LSR:, and content, which occur if invalid data is provided in a field. A rejected service request must be corrected and re-submitted before provisioning can begin.
Repeat Report	Any trouble report that is a second (or greater) report on the same telephone number/circuit ID and at the same premises Address within 30 days. The original report can be any category, including excluded reports, and can carry any disposition code.
Service Group Type	The designation used to identify a category of similar services, .e.g., UNE loops
Service Order	The work order created and distributed in ILECs systems and to ILEC work groups in response to a complete, valid service request.
Service Order Type	The designation used to identify the major types of provisioning activities associated with a service request
Service Request	The transaction sent from the CLEC to the ILEC to order services or to request a change(s) be made to existing services.
Standard Interval	The interval that the ILEC quotes to its customers with respect to how long it will take to provision a service request. These intervals are standardized by specific service type and type of service modification requested ILECs publish these standard intervals in documents used by their own service representatives as well as ordering instructions provided to CLECs. POTS services do not have standard intervals;, their installation intervals are based on force available and workload. They may change as frequently as twice a day.

DEFINITION OF TERMS

TERM	DEFINITION
Subsequent Reports	A trouble report that is taken on a previously reported trouble prior to the date and time the initial report has a status of "cleared".
Summarized Charges	Billing charges that are aggregated on the bill, rather than individually itemized, e.g., local usage minutes on resale or retail calls, which are listed on the bill as "xx" minutes with no call detail.
Tandem Switch	Switch used to connect and switch trunk circuits between and among Central Office switches.
Time to Restore	The time interval from the receipt, by the ILEC, of a trouble report on a customer's service to the time service is fully restored to the customer.
To Be Called Cut	A type of coordinated customer conversion, which involves the CLEC calling the ILEC to signal the ILEC that it should start the customer conversion. (Pacific Bell term)
Trouble Cause Code	A code identifying the known or suspected cause of a trouble condition.
Trouble Disposition	A code identifying the end result of diagnostic and/or repair activities on a customer trouble report.
Usage Data	Data generated in network nodes to identify switched call data on a detailed or summarized basis. Usage data is used to create customer invoices for the calls.
Usage Records	The individual call records created in a switch to report the date, time, duration, calling and called numbers associated with a given call
Virtual Collocation	Shall have the meaning set forth in 47 C.F.R. Section 51.5.

CALIFORNIA OSS OII
PERFORMANCE MEASURES: GLOSSARY OF ACRONYMS

ACRONYM	DESCRIPTION
ADSL	Asymmetric Digital Subscriber Line
ALI	Automatic Line Information (for 911/E911 systems)
AS	Affecting Service (type of trouble condition)
ASI	Advanced Services Inc. (data subsidiary of SBC)
ATIS	Alliance For Telecommunications Industry Solutions
BDT	Billing Data Tape
BOS	Billing Output Specifications
BRI	Basic Rate Interface (type of ISDN service)
CABS	Carrier Access Billing System
CARE	Customer Repair Center (GTE)
CBSS	Customer Billing Service System (GTE)
CESAR	Carrier Enhanced System for Access Request
CHC	Coordinated “Hot” Cut
CKT	Circuit
CLEC	Competitive Local Exchange Carrier
CO	Central Office
CORBA	Common Object Request Broker Architecture (Pre-ordering standard)
CPE	Customer Premises Equipment
CPUC	California Public Utilities Commission
CRIS	Customer Record Information System
CSB	Customer Service Bureau (PB retail repair center)
CSR	Customer Service Record
DA	Directory Assistance
dB	Decibel
DID	Direct Inward Dialing
DS0	Digital Service 0
DS1	Digital Service 1
DS3	Digital Service 3
E911 MS	E911 Management System
EAS	Equal Access Service
EDI	Electronic Data Interchange
EMI	Exchange Message Interface
EUCL	End User Carrier Line charge
FDT	Frame Due Time
FOC	Firm Order Confirmation
GTE	General Telephone Company
GTT	Global Title Translations
GUI	Graphical User Interface
HDSL	High-bit-rate Digital Subscriber Line
HICAP	High Capacity Digital Service
IEC	Inter-exchange Carrier
ILEC	Incumbent Local Exchange Carrier
I, N, T, C, M	Service Order Types - I (install-GTE), N(new-PB), T(to or transfer-PB), C(change)and M(move-GTE)
ISDN	Integrated Services Digital Network
IW	Inside Wire
LATA	Local Access Transport Area
LERG	Local Exchange Routing Guide

CALIFORNIA OSS OII
PERFORMANCE MEASURES: GLOSSARY OF ACRONYMS

ACRONYM	DESCRIPTION
LNP	Local (or Long Term) Number Portability
LOC	Local Operations Center (PB repair and coordination center for CLEC activity)
LSC	Local Service Confirmation or Local Service Center (PB)
LSMS	Local Service Management System
LSR	Local Service Request
MAC	Missed Appointment Code
NDM	Network Data Mover
NOMC	National Open Market Center (GTE)
NPAC	Number Portability Administration Center
NXX	Telephone number prefix
OBF	Ordering and Billing Forum
OOS	Out of service (type of trouble condition)
OSS	Operations Support System
PB	Pacific Bell
PBX	Private Branch Exchange
PICC	Primary Interexchange Carrier Charges
PNP	Permanent Number Portability (same as LNP)
PON	Purchase Order Number
POTS	Plain Old Telephone Service
PRI	Primary Rate Interface (type of ISDN service)
SBC	Southwestern Bell Corporation
SCP	Service Control Point
SDA	Separate Data Subsidiary
SGT	Service Group Type
SORD	Service Order Retrieval and Distribution (PB service order creation system)
SOT	Service Order Type
SS7	Signaling System 7
STP	Signaling Transfer Point
TBCC	To Be Called Cut (PB)
TN	Telephone Number
UNE	Unbundled Network Element
VGPL	Voice Grade Private Line
xDSL	(x) Digital Subscriber Line

MISSED APPOINTMENT CODES – PACIFIC BELL
MAC – COMPANY REASONS

CB	Marketing Error. LSC/ Business Office gave wrong due date or ordered incorrect product/service
CO91	No Access to Terminal Or Protector
CO92	No Electrical Permit-Company
CO93	All Other Company Reasons (Tone Back)
CO94	Joint Marketing Contractor
CO95	Civil Unrest, No Access
CO96	National 800 database to Facilities
CO97	Malfunction of Mechanized Service Order Systems i.e. SORD, COSMOS, FACS, MARCH, PBOD
CO98	NFWK Service Order Sent To Field and Due Date Missed
CO99	Missed Appointment Window - Senate Bill 101 (System Failure)

COMPANY WORK LOAD

CL71	Installation-Force/Load Imbalance
CL72	Weather Conditions
CL73	Sanctioned Work Stoppage Against Pacific Bell
CL74	Emergency Conditions, Earthquakes, Floods
CL75	800 Service Center Work Load Imbalance
CL79	Missed Appointment Window - Senate Bill 101 (Work Load)

EQUIPMENT SUPPLY

CE81	Lack of Normally Ordered Facility Equipment or Supplies
CE82	Lack of Specially Ordered Facility Equipment or Supplies
CE83	Other Facility Equipment Problems

COMPANY FACILITIES

CF61	Lack of Outside Plant
CF62	Lack of C/O Facilities
CF63	BSW
CA	Lack of Assignment
CS	Switching Error

MISSED APPOINTMENT CODES – PACIFIC BELL

MAC – CUSTOMER REASONS

NO ACCESS	DESCRIPTION
SA01	None on Prem Left Notice
SA02	Agent/Mgr Not On Prem Left Notice
SA03	Denied Access To Term. On Cust. Prem Left Notice
SA04	Manager Refused Access Left Notice
SA05	Manager Had No Key Left Notice
SA06	Security Type Building
SA07	Unable to Locate Other Designated Party
SA08	Dog/Other Safety Hazard On Premises
SA09	No Response To Call Before Going Number (3 Or More Attempts Made)
SR20	Subscriber In Independent Company No Facility In Independent Company
SR21	No Pole
SR22	No Conduit
SR23	Conduit Plugged
SR24	inc. Full No Spares, Referred to Building Owner, No Authorization./Pre-Authorization to Repair
SR25	No Trench
SR26	Not Authorized To Sign Labor Receipt
SR27	Customer Requests Later Due Date From Tech.
SR28	Building Not Ready
SR29	Electric Power Not Available

CUSTOMER REQUESTS LATER DUE DATES

SL31	Customer Called Company before Tech. Arrived
SL32	Pre-Survey Contact Customer Requests Changing of Due Date

ALL OTHER CUSTOMER REASONS

SO41	Minor Daily Access
SO42	Customer Requested Additional Work
SO43	Customer Gave Wrong Address
SO44	Access Refused
SO45	Access Didn't Know Installation Locations
SO46	Mgr./Owner OK Needed For Exposed Wiring
SO47	Mgr./Owner OK Needed To Drill Hole
SO48	Customer Required To Pay Deposit
SO49	Missed Appointment Window- Senate Bill 101 (Customer Gave Wrong Address)
SO50	Vendor Problem Regarding CPE Term Equipment Either Not Delivered/Installed or Removed

JEOPARDY
MISSED APPOINTMENT CODES -GTE

Standard OBF Jeopardy Code	**Description**
1A	Inter Office Facility Shortage
1B	Scheduling/Work Load
1C	Customer Not Ready
1D	No Loop Available
1E	End User Not Ready
1F	Provider Missed Appointment
1G	No Access to End User Premise
1H	Central Office Freeze
1J	Special Construction
1K	Natural Disaster (Flood, etc.)
1L	Frame Due Time Cannot Be Met
1M	Requested Due Date Is Not Available
1N	Due Date and Frame Due Time Cannot Be Met
1P	Other
1Q	Assignment Problem
1R	Customer Could Not Be Reached at the Can Be Reached Number (CBR)
1S	Building Not Ready, Customer Will Advise
1T	Pole At Site Not Set
1W	Entrance Facilities Required
1X	Not Technically Feasible
1Y	No Central Office Equipment Available
1Z	Other Local Exchange Company Not Ready
2A	CLEC order request error
2B	Work order pending

Verizon has adopted standard OBF jeopardy codes, listed above.

DISPOSITION CODES

	PACIFIC BELL		GTE
01	TERMINAL EQUIPMENT	04	NETWORK FACILITIES
02	COMMUNICATIONS EQUIPMENT	05	COIN/COINLESS
02	OTHER STATION EQUIPMENT	05	E911
02	TERMINAL EQUIPMENT	06	OUTSIDE PLANT
03	NETWORK TERMINATING FACILITIES	07	INTEROFFICE FACILITIES
04	OUTSIDE PLANT	09	SERVICE ORDER
05	CENTRAL OFFICE	10	RECORDS
06	CUSTOMER MISUSE	11	CARRIER (FIELD) OR CONCENCENTRATOR
07	TEST OK	12	CENTRAL OFFICE
08	FOUND OK - IN	13	TEST OKAY
09	FOUND OK – OUT	15	CAME CLEAR
10	REFERRED OUT	16	CUSTOMER
12	NON-TELCO PROVIDED	17	EXCLUDE
13	INTER-EXCHANGE CARRIER/INDEPENDENT COMPANY	18	REFERRED OUT
		19	CPE
	PACIFIC BELL CAUSE CODES		
1	TELCO EMPLOYEE		
2	NON-EMPLOYEE		
3	PLANT OR EQUIPMENT		
4	WEATHER		
5	OTHER		
6	UNKNOWN		

Decision 01-01-037 January 18, 2001

BEFORE THE PUBLIC UTILITIES COMMISSION OF THE STATE OF CALIFORNIA

Order Instituting Rulemaking on the Commission's Own Motion into Monitoring Performance of Operations Support Systems.	Rulemaking 97-10-016 (Filed October 9, 1997)
Order Instituting Investigation on the Commission's Own Motion into Monitoring Performance of Operations Support Systems.	Investigation 97-10-017 (Filed October 9, 1997)

INTERIM OPINION ON PERFORMANCE INCENTIVES

TABLE OF CONTENTS

TABLE OF CONTENTS

TITLE **PAGE**

TABLE OF CONTENTS

Summary

The Telecommunications Act of 1996 (TA96 or the Act) was a major step in the process of opening previously monopolistic local telephone service markets to competition. To foster competition, the act requires the incumbent local exchange carriers (ILECs) to provide competing carriers access to any necessary ILEC infrastructure, including the incumbents' operations support systems (OSS). OSS includes pre-ordering, ordering, provisioning, maintenance, billing, and other functions necessary to providing various telephony services. For competition to occur, the competitive local exchange carriers (CLECs) must be able to access these services in the same manner as the ILEC.

For example, for pre-ordering, a CLEC must be able to access customer information relevant to the service being ordered, so that the CLEC can tell its customers what options they have. For ordering, a CLEC needs to be sure that the ordering process for its customers takes no more time than for ILEC customers. Similarly, for provisioning, a CLEC needs to be sure that the time the ILEC takes to actually install or provide a new telephone service for CLEC customers is no longer than for ILEC customers. Delays or inaccuracies in these and the other OSS functions could discourage potential customers from doing business with the competitors.

Under its authority to implement the Act, the Federal Communications Commission (FCC) has strongly encouraged that regulatory remedies be established to ensure ILEC OSS performance does not present barriers to competition. While not an outright prerequisite for FCC approval of Regional Bell Operating Companies' (RBOC) applications to provide in-region interLATA service under § 271, the FCC has indicated that such applications must be in the public interest. In its evaluation of the public interest, the FCC states that, "the

fact that a BOC will be subject to performance monitoring and enforcement mechanisms would constitute probative evidence that the BOC will continue to meet its section 271 obligations and that its entry would be consistent with the public interest."[1] As a consequence, we will establish a performance remedies plan to identify and prevent or remove any barriers. The three critical steps for any performance remedies plan are performance measurement, performance assessment, and the corrective actions necessary if performance is deemed harmful to competition.

The California Public Utilities Commission (Commission or CPUC) has established performance measures in a parallel proceeding in this docket. Our decision today establishes an interim performance assessment plan. We have created a set of procedures for assessing the performance measurement results to identify competitive barriers. In effect, we have set forth a self-executing decision model that applies barrier-identifying criteria to the performance measurement results. A self-executing plan is one that requires no further review and no new proceedings. Explicit, objective, data-based standards are established that automatically calculate and determine the existence of "competitive barrier" performance. Statistical tests identify barriers when ILEC performance to its own customers can be compared to ILEC performance to CLEC customers. Explicit performance levels, called benchmarks, identify barriers when there is no comparable ILEC performance.

This decision model now enables us to proceed to the final step of the remedies plan, establishing the incentives that will be tied to any deficient

---

[1] *Bell Atlantic New York Order* ("FCC BANY Order"), 15 FCC Rcd at 3971, ¶ 429.

performance identified by the model. The overall goal of the plan will be to ensure compliance with the FCC's directive that OSS performance shall provide competitors a true opportunity to compete.

Background

On October 9, 1997, the Commission instituted this formal rulemaking proceeding and investigation to achieve several goals regarding Pacific Bell's (Pacific) and Verizon California, Inc.'s (Verizon CA)[2] OSS infrastructure. One objective of this docket (the OSS OII/OIR) is to assess the best and fastest method of ensuring compliance if the respective OSS of the ILECs do not show improvement in implementation or meet determined standards of performance. Another related objective is to provide appropriate compliance incentives under Section 271 of TA96, which applies solely to Pacific[3], for the prompt achievement of OSS improvements.

To further these specific objectives, the ILECs and a number of interested CLECs participated in a series of meetings jointly conducted through the OSS OII/OIR proceeding and the 271 collaborative process[4]. In October 1998, a group

---

[2] Verizon CA was previously named GTE California Incorporated. Hereafter, Pacific and Verizon CA will be referred to collectively, as the ILECs.

[3] As a Bell Operating Company (BOC), Section 271 specifically applies to Pacific.

[4] From July through mid-August 1998, Pacific, AT&T Communications of California Inc. (AT&T), MCI WorldCom (MCI W), Sprint Communications, Electric Lightwave, Inc., ICG Telecom Group, Inc., Covad Communications (Covad), MediaOne Telecommunications of California, Inc., Cox California Telecom, LLC, Northpoint Communications, California Cable Television Association, and staff entered into a collaborative process and jointly worked on developing solutions to the flaws in Pacific's 1998 draft 271 application. Verizon CA observed one collaborative meeting on penalties, but otherwise did not participate. (Verizon CA Response to Motion to Accept

Footnote continued on next page

of the interested parties filed joint comments setting forth their various positions on the issues discussed during the meetings. Following a pre-workshop conference in January 1999, the assigned Administrative Law Judge (ALJ) and the Telecommunications Division staff (staff) convened a 7-day technical workshop[5] on the respective performance incentive plans of Pacific and the participating CLECs. Pacific and the CLECs filed concurrent opening briefs on March 22, 1999, and concurrent reply briefs on April 5, 1999.

Pursuant to ALJ Ruling, Verizon CA filed its proposal on incentives for compliance with performance measures on May 3, 1999. The CLECs responded to the proposal on May 11, 1999. On July 12-14, 1999, the ALJ and staff convened a technical workshop on Verizon CA's performance incentive plan in relation to the CLECs' plan[6]. The parties filed concurrent opening briefs on July 28, 1999, and concurrent reply briefs on August 4, 1999. On August 12, 1999, Verizon CA petitioned to have submission set aside and supplemental comments accepted. The CLECs responded to the petition on August 27, 1999.

On November 22, 1999, the assigned Commissioner noted in a ruling (the ACR) that staff and its technical consultants had advised him that the performance incentive plans that the parties had submitted were significantly flawed. The ACR set forth the framework of a performance remedies plan that it encouraged Pacific, Verizon CA and the CLECs to analyze and comment upon

---

Joint Comments regarding Report on Performance Incentives, footnote 2 at 2 (October 20, 1998)).

[5] February 5, 8-11, and 23-24, 1999.

[6] The CLECs submitted their plan in both the Pacific and Verizon CA portions of the proceeding.

with the overall goal of developing a common and acceptable approach to implementing the performance plan. The parties filed opening comments on the ACR on January 7, 2000. Pacific and the Office of Ratepayer Advocates[7] (ORA) included new performance incentive plan proposals with their initial comments. The parties filed reply comments on January 28, 2000.

On March 27, March 28 and March 30, 2000, the ALJ, assisted by staff, convened a facilitated workshop that focused exclusively on the performance assessment part of three performance remedies proposals: (1) the ACR-proposed plan; (2) the new Pacific plan, and (3) the ORA plan. The parties submitted opening and closing briefs on April 28 and May 5, 2000, respectively.

Performance Remedies Plan Fundamentals

The TA96[8] and the FCC's implementing rules require Pacific and Verizon CA to provide CLECs with nondiscriminatory access to unbundled network elements (UNEs), including OSS. The FCC commented generally that ILECs must provide the CLECs with access to the pre-ordering, ordering, provisioning, billing, repair, and maintenance OSS sub-functions pursuant to the Act such that the CLECs are able to perform such OSS sub-functions in "substantially the same time and manner"[9] as the ILECs can for themselves.

---

[7] ORA had monitored this phase of the OSS OII prior to its January 7th submission.

[8] Section 251(c)(3).

[9] *Implementation of the Local Competition Provisions of the Telecommunications Act of 1996, CC Docket No. 96-98*, First Report and Order, 11 FCC Rcd, at 15763-64 (1996) (Local Competition First Report and Order).

The Act does not expressly mandate the establishment of either performance measures or incentives, though the FCC has stated that the most probative evidence that the CLECs are provided with nondiscriminatory access to OSS will be evidence of actual commercial usage evaluated under a set of Commission-approved performance measures. Similarly, TA96 and the implementing rules have no stated requirement for an additional customer economic effect test. The FCC has stated that an ILEC may demonstrate statistically that the differences in measured performance are the result of random variation in the data, as opposed to underlying differences in behavior. The phrase "underlying differences in behavior" means differences in the statistical distributions of the ILEC and the CLEC that are generating the performance outcomes.[10] Thus, equality of distributions (when the ILECs' and the CLECs' distributions are the same) is a sufficient condition for parity according to the FCC.[11]

The cornerstone of any performance incentive structure is how parity is defined, since it is on those occasions when the ILECs are out of parity that

---

[10] Roughly speaking, distributions are different when average performance and range of performance (variability, distribution) are different. For example, CLEC customer phone service provisioning could take 7 days on the average, whereas ILEC customer service provisioning could take 6 days. In this example, average performance for the ILEC is better than for the CLEC by one day. For variability, even with equal ILEC and CLEC averages of 7 days, CLEC provisioning times could range between 1 and 13 days, whereas ILEC averages could range between 6 and 8 days. In this example, performance for the ILEC is less variable, and thus more predictable. ILEC customers could be told that their new service would be installed in 8 days or less, in contrast to CLEC customers who could only be told that their service would be installed in 13 days or less.

[11] Id.

incentive payments will be made. This Commission's definition of parity generally incorporates the above-stated objectives of the TA96 and the FCC. Thus, parity means that the ILEC is providing services in substantially the same period of time and manner (including quality) to the CLECs as it is providing to itself. Further, it will be helpful to rely on statistical testing and benchmarks to infer whether or not parity has been achieved. Consequently, we endeavor to ensure that the CLECs have OSS access that is at least equal to the ILECs' own access.

Initial Proposed Plans

This section provides an overview of the history of this proceeding, and focuses on the parties' various positions and plans. Brief explanations of statistical concepts are presented with the limited purpose of identifying parties' positions. A more detailed explanation of statistical concepts accompanies our deliberation in the section titled "Selection of the Statistical Model."[12]

Plan Principles

Pacific initially developed a statistical approach to determining compliance with TA96's nondiscriminatory access standard structured on three central principles. First, the remedy plan must not impose payments on Pacific when nondiscriminatory or parity treatment is provided.[13] However, Pacific conceded that, given the nature of the statistical models applied, it was difficult to drive the parity payment amount closer to zero without lowering the out-of-parity

---

[12] Readers wishing elementary or more detailed statistical explanations before reading this section may wish to first read the section titled: "Selection of the Statistical Model."

[13] "The expected cost for parity treatment should be zero."

payments substantially. (Pacific's 1999 Opening Brief on Performance Remedies at 2-3.)

Second, if Pacific does not provide parity treatment, then payment amounts to the CLEC should have some reasonable relationship to the level of performance provided.[14] Pacific argued that remedy amounts should not be enormous when the level of performance deviates from parity by only small amounts or in isolated incidents. Thus, the levels of remedies should start relatively low and increase commensurately with the level of nonperformance. Id. at 3.

Third, remedy payments should motivate Pacific to provide nondiscriminatory service, but should not motivate the CLECs to favor receiving large remedy payments.[15] Therefore, the remedy amounts must not be so high that a CLEC would be more desirous of receiving poor service and collecting large payments than receiving nondiscriminatory service. Id.

The CLECs also based their initial incentive proposal on three principals. They declared that the incentives must be in an amount sufficient to cause Pacific to meet its parity obligations. Second, the incentives must be self-executing without broad opportunity for circumvention or lengthy delay in the payment of the consequences. Finally, the CLECs asserted that the structure of the plan must be fairly simple to implement and monitor.

---

[14] "Payments should bear a reasonable relationship to level of performance."

[15] "CLECs should not be motivated to receive large remedy payments."

Parity and Statistical Model Elements

In its initial performance incentive proposal, Pacific defines parity to mean delivering services to CLEC customers from the same processes as delivered to ILEC customers. When organizationally it is not possible to have the same processes, Pacific then defines parity to mean that the ILEC must deliver services with the same properties to the CLEC as delivered to the ILEC. The definition for parity, and the test for parity, appears to be the same, i.e., 1.645 standard deviations from the mean.[16] (Pacific 1999 Opening Brief at 5-6 and 13-15.)

Verizon CA contends that parity only requires that CLEC ordering processes be performed in "substantially the same time and manner" as the ILEC's like processes. It claims that ILECs have unavoidable variations in their own processes, and as long as the ILEC and CLEC distributions are substantially the same, parity is present. Verizon CA also considers the appropriate test for parity to be average performance within 1.645 standard deviations of the mean. (Verizon CA 1999 Opening Brief at 5.)

The CLECs define parity as equal service for the ILEC and the CLEC. The CLECs want zero (0) standard deviations from the mean for the definition of

---

[16] A standard deviation is a standardized statistic measuring how dispersed scores are. A low standard deviation indicates scores are grouped closer to the mean than scores with a higher standard deviation. When applied to a normal or "bell-shaped" curve, the standard deviation provides helpful information about the dispersion of scores: 68.3 percent of all scores lie within one standard deviation of the mean (plus or minus one standard deviation, 95.4 percent lie within 2 standard deviations, 99.7 lie within 3 standard deviations, and so forth. In the present application, 1.645 standard deviations above the mean encompass 95 percent of the scores. So under conditions of random selection, a score greater than 1.645 standard deviation would be selected 5 percent or less of the time.

parity, but have offered that a test for determining parity could be one (1) standard deviation from the mean. (CLECs' 1999 Opening Brief at 4-15.)

In its May 3, 1999 preliminary statement, Verizon CA embraced each of the core principles Pacific and the CLECs set forth, and asserted that the concepts need not be mutually exclusive. Moreover, it added the following seven principles of its own to the "ideal" incentive plan. First, a design objective of the plan should be that no incentive payments should be made when parity exists. Consequences should be economically significant, not just statistically significant. Further, the incentive structure should provide that the incentive payment equals the resource cost of meeting the standard. Regular review periods are necessary. The incentive mechanism should not result in large administrative costs. There must be some "off-ramps" in a self-executing incentive system to deal with certain circumstances. Finally, with an eye to the future, the plan should be symmetrical across all parties. (Verizon CA Brief on OSS Performance Incentives at 2-5.)

Test for Determining Compliance with Parity

Pacific originally proposed using a standard Z-test[17] for purposes of determining compliance with parity. The CLECs objected to the standard Z-test,

---

[17] **Standard Z-test** : Z = Difference / Standard deviation of the difference
Where: Difference = Pacific Average – CLEC Average.

Standard deviation of the difference = Square root of ((Variance of Pacific x 1 / Pacific sample size) + (Variance of CLEC x 1 / CLEC sample size)).

Or, assuming the variances for Pacific and the CLEC are equal, the variances are pooled together: Standard deviation of the difference = Square root of ((Pooled variance of Pacific and CLEC samples) x (1 / Pacific sample size + 1 / CLEC sample size)).

which utilizes the individual variances of the Pacific and CLEC samples, arguing that Pacific could manipulate the variance of the CLEC sample. Pacific responded that the standard Z- test was adequate because any alleged manipulation of the CLEC sample variance would be readily apparent.

The CLECs speculated that Pacific could increase the variance of the CLEC sample, which would reduce the probability that Pacific would be found out-of-parity.[18] In response, they proposed the "Modified Z-test,"[19] which modifies the standard Z-test by using only Pacific's sample variance. In the "spirit of collaboration," Pacific offered to use the CLECs' proposed Modified Z-test on a trial basis, and then test it in order to evaluate whether the Modified Z-test yielded "fair and accurate results." Verizon CA agreed to use the Modified Z-test to assess parity subject to review and modification following a six-month interim implementation period.

---

[18] An increased CLEC variance theoretically could increase the size of the Z-test denominator without affecting the numerator, thus reducing the resulting Z-test statistic and reducing the chances of identifying out-of-parity situations.

[19] **Modified Z-test:** Z = Same as Z-test.

Where:

Difference = Same as Z-test.

Standard deviation of the difference =

Square root of (Variance of Pacific x (1 / Pacific sample size + 1 / CLEC sample size)).

Minimum Sample Size

Pacific initially desired a minimum sample size of thirty occurrences.[20] In the "spirit of cooperation," Pacific was willing to lower the sample size to twenty, with the caveat that the impact of smaller sample sizes be evaluated during a review period in the not too distant future. Pacific also accepted benchmark measures for a specific list of rare submeasures.[21] That is, parity measures with rarely occurring activity were essentially to be converted to benchmark measures.

The CLECs acknowledge that many of their number will have fewer than thirty observations (e.g., orders) in a month for some measures. They want to ensure that a requirement of a larger sample size does not passively provide an acceptable level of performance to the ILEC. Therefore, the CLECs preferred sample sizes as small as one, but suggested a minimum sample size of five for parity submeasures. The CLECs also accepted the benchmark measures for the specific list of rare submeasures.

Verizon CA supported the use of "table lookup"[22] for sample sizes exceeding 50 CLEC transactions. Noting that there is a lack of experience using

---

[20] A sample size of thirty is a standard textbook "rule-of-thumb" sample size cutoff for parametric statistical testing such that distributional assumptions can be anticipated to be met for most situations.

[21] A "measure" defines how performance will be measured for a specific OSS function, such as ordering, across several service types, such as residential telephone service, business telephone service, DSL service, etc. A "submeasure" applies the specified "measure" methods to individual service types, for example, either residential telephone service, or business telephone service, or DSL service, etc

[22] The statistical test produces a test value. The test value can then be "looked up" in a table to determine statistical significance. In most cases a normal approximation or a "t"

Footnote continued on next page

the Modified "t" statistic[23] for non-normal samples, Verizon CA advocated using permutation tests for sample sizes between 20 and 50. (Verizon CA 1999 Opening Brief at 33-34.) For sample sizes less than 20, Verizon CA originally proposed that the CLECs and it should explore, during the interim development period, use of: (1) permutation tests; (2) aggregation of results across sub-measures; (3) aggregation of results across CLECs; and (4) possible exclusion of a given measure from performance incentive assessment. During the interim period, Verizon CA stated that it would also rely, to the extent practicable, on "exact methods"[24] to determine achieved significant levels for small sample tests on proportions. (Id. at 34.)

Alpha Level/Critical Value

Pacific and Verizon CA proposed a Z statistic of greater than 1.645 standard deviations (critical value) to determine “out-of-parity.” A 1.645 standard deviation corresponds to a five percent (one-tailed) Type I error, or “alpha.” A Type I error is rejecting the null hypothesis (i.e., parity service)[25] when it should not be rejected. A Type II error is accepting the null hypothesis when it should not be accepted. “Alpha” is the probability of a Type I error and “beta” is the probability of a Type II error. Values of 1, 5, and 10 percent alpha levels are the most common “textbook” values.

---

distribution table is used to determine the Z or t statistic that must be exceeded for a performance failure finding.

[23] The “Modified t-test” is a variant of the Modified Z-test used for sampling distributions of small sample mean, as discussed later in this Decision.

[24] The term "exact methods" is defined as performing all possible permutations.

[25] A “null hypothesis” proposes that there are no differences between the true means.

The null hypothesis in this application poses that ILEC and CLEC performance are in parity. A Type I error is identifying the ILEC as not providing parity service (i.e., the ILEC is providing worse service to CLECs than to itself) when in fact the ILEC is providing parity service. A Type II error is identifying the ILEC as providing parity service when in fact it is not providing parity service. Pacific wanted to be limited to a five-percent probability of being identified as not providing parity service when in fact it is providing parity service.

The CLECs recommended an equal error methodology be employed for setting the errors. This essentially calculates and equates the Type I and Type II errors for each submeasure each month. The CLECs ultimately suggested that a Z statistic of greater than 1.04 standard deviations (critical value) should identify "out-of-parity" conditions. A 1.04 standard deviation corresponds to a fifteen percent (one-tailed) Type I alpha level. The CLECs were concerned with Type II errors, not just Type I errors. By making the critical alpha level larger, the CLECs worried less about the beta error.[26] Thus, the CLECs wanted at least a fifteen-percent probability limit for identifying Pacific and Verizon CA as not providing parity service when in fact they are providing parity service, because they believed that this would correspond more closely to an equal probability of identifying non-parity service as parity service.

Assigned Commissioner's Ruling and Proposed Plan

By ruling issued November 22, 1999, the assigned Commissioner assessed the submitted proposed plans and set forth his concerns about them (the ACR).

---

[26] As the critical alpha level is increased (e.g., from 0.05 to 0.15), beta decreases.

The ACR noted that the existent ILEC models and the CLECs' model appeared distinct and incompatible. In addition, the parties revealed considerable misunderstanding and confusion about the two sets of respective model assumptions and calculations. It was difficult to sort out the relative impacts of each of the respective components of the two differing model approaches. Moreover, the end result outcomes of the two models were highly uncertain because both the modeling approaches were trying simultaneously to design and implement the total model (both the performance assessment model elements and the incentive plan elements) without the benefit of an implementation and data calibration structure.

While the plans' proponents had articulated numerous core concepts, no distilled set of principles supported both plans. There also appeared to be little rationale for the incentive levels implicit in either plan. It is unlikely that either plan could be implemented as designed. Moreover, both models might impose costs when evidence suggests parity service, and both models might not impose costs when evidence suggests non-parity service. During the February 1999 technical workshop, each proposed plan produced dramatically different payments due to different input assumptions. Both plans were also very sensitive to minor changes in assumptions. These problems were not due to an attempt to keep the plans simple; both the ILECs' and CLECs' plans were very complex. Accordingly, we affirm the ACR's evaluation of the initial ILECs' and CLECs' plans.

The ACR expressed the need to have one common interim model framework of analyses for review and discussion, and for use by all concerned parties in order to implement the performance remedies plan. One interim performance remedies plan model and set of explicit assumptions, would allow common quantitative analyses to be performed and estimates to be developed. All key model assumptions would be explicit, and the policy ramifications of these assumptions would be clear.

The ACR proposed that a common and feasible approach to implement the necessary performance remedies plan[27] be developed with the assistance of the ILECs and the CLECs. It noted that to achieve the single common model framework, there needed to be an unwinding of the performance assessment model elements and the incentive plan elements that the parties merged together from the outset. To that end, the ACR proposed an initial conceptual performance measurement statistical model, and asked the parties to respond to specific questions about the model. Further, it proposed that the Commission implement a fully functioning, self-executing performance remedies plan during a six-month pilot test period.

We concur with the ACR assessment that a single model approach would allow the Commission to make informed policy decisions about the performance remedies plan. A single model approach focuses on the goal of parity service by the ILECs, economic incentives paid by the ILECs, and / or a change in ILECs' operations support to the CLECs. The end goal is certainly not just to have complex statistical measurement theory applications. There may be a variety of statistical measurement approaches that can all achieve the same basic economic and operations incentives by using different incentive plan structures and amounts, in combination with different measurement approaches.

---

[27] To avoid confusion with the work going on in the Performance Measurement segment of this proceeding, what is essentially the "performance measurement, assessment and incentive" plan will be referred to as the "performance remedies" plan.

A single common interim model and a single set of explicit assumptions should allow calibration of end result economic outcomes both before and after a six-month pilot test period using actual empirical data. The interim pilot test period can assist the Commission in determining the appropriate levels of long-term economic incentives. Long-term incentive impacts can be calibrated in relation to one model, one common set of assumptions, and actual test period empirical data. Penalty amounts and structures can still be set and paid during the pilot test period, and they can be applicable only during this interim period, unless otherwise determined.

ACR Plan Statistical Model Elements

Noting the ILECs' and CLECs' distinct views on standard and Modified Z-tests, the ACR questioned whether there would be a way to determine if the Modified Z-test yields "fair and accurate results." Of interest are differences in the results if the standard Z-test was used rather than the Modified Z-test. Such differences would be due to disparities between the variances of Pacific and the CLECs. Regarding the CLEC position that the variance of the CLEC sample could be potentially manipulated, the ACR stated that concern about the possibility of manipulation should not direct the test procedure.

The ACR suggested that the optimal course might be for the Commission to proceed with the standard Z-test on a trial basis to be evaluated after a six-month test period. The proposed Modified Z-test[28] applies an experimental

[28] It also holds the possibility of manipulation.

argument[29] to an observational situation. There are no other academic precedents for our application of this particular modified calculation. The ACR stated that it was doubtful at this point whether any further complicating modifications to the statistical methodology for determining compliance with parity would be worth the benefits without first trying the standard Z-test.

The standard Z-test is the most common method to compare two population means, under the following key assumptions:

1. Underlying distributions are not too skewed (i.e., they are not too different from a normal bell shaped curve).
2. Sample sizes are reasonably large.
3. Observations are independent measurements from the same processes (e.g., phone service installation operations).

If the variances are known to be equal, then a pooled, or common, variance estimate is used. If the variances are known to be unequal, then both separate variances are used. If it is unknown, *a priori*, whether the population variances are equal or not, then an initial test compares the variances. Based on this first test, either the separate or pooled variance estimate is used.

The genesis of the Modified Z-test assumes the contention that Pacific could manipulate the variance of the CLEC sample. While such manipulation might be possible, it seems equally likely that Pacific could simultaneously manipulate the mean of the CLEC sample, and the variance and mean of the corresponding Pacific sample. The ACR proposed to first test for variance equality between Pacific and CLEC results. If the variances prove to be unequal,

---

29 Brownie, Cavell, Boos, D., and Hughes-Oliver, J. *Modifying the t and ANOVA F Test When Treatment Is Expected to Increase Variability Relative 2 Controls,* 46 Biometrics at 259-266 (1990).

the ACR suggested that it might be necessary to use the standard Z-test with both variances. In either case, parity will be assumed to exist when the differences in the measured results for both the ILECs and the CLECs in a single month, for the same measurements, are less than the critical value[30] of the Z-test.

Early on, the CLECs implied that the difference between the standard Z-test and the Modified Z-test could measure Pacific's ability to manipulate the data. Since both Pacific and the CLECs have agreed to use the Modified Z-test during a pilot test period, the ACR raised the possibility that both the standard and Modified Z-tests might be calculated and evaluated over the six-month pilot test period. However, the ACR further proposed that if both tests were run, actual calculations during the trial test period would be based on the standard Z-test. The results of the evaluation might suggest that the decision as to which form of Z-test to use might be moot, since all choices might identify the same situations as being out-of-parity.

The ACR also suggested that during the six-month pilot test period, sample distributions could be reviewed to explore whether the distributions meet the above-stated underlying assumptions of the Z-test. At the end of this six-month pilot test period, there could be a reconsideration of whether any variety of Z-test should be used, or whether nonparametric tests[31] might be more appropriate. All of the Z-tests described by Pacific and the CLECs are parametric

---

[30] The critical value of the Z-statistic corresponds to a critical alpha value. The rejection region encompasses the critical Z-statistic and larger Z-statistic values, which correspond to critical alpha and smaller alpha values.

[31] Distribution-free tests based on medians or ranks; that is, tests not dependent on assumptions about distributions, such as normality.

tests. They assume observations are independent and are generated from the same process with a relatively well-behaved distribution.[32] However, the ACR questioned the independence of the observations and the shapes of the distributions, especially the CLEC distributions. The ACR suggested that if these characterizations were accurate, over the long-term it might be better to use nonparametric tests.

Finally, the ACR noted that there appeared to be some confusion regarding the concept of samples versus entire populations. If, as the ACR surmised, it would be appropriate to assume we had the entire population of measurements during a time period, as with production output, then it might make sense to ultimately utilize concepts of statistical process control to monitor and modify the procedures when they appear to have gone, or likely will be going, out of control. For example, a production monitoring and control methodology[33] could utilize the mean and variance of the ILEC (essentially as a benchmark against which CLEC measurements are compared). This could be performed using a Z-test-based chart set only on the mean of CLEC measurements against the historic mean and variance[34] or other statistics of the ILEC. Or similarly, a permutation test could be used.

---

[32] "Well-behaved" refers to distributions where a resulting distribution of sample means is not deviant enough from a normal distribution to cause inaccuracies – discussed later in this decision.

[33] For example, a Shewart Quality Control chart. R. Mason, R. Gunst, and J. Hess, Statistical Design And Analysis Of Experiments With Applications To Engineering And Science at 65 (1946).

[34] Or cumulative values.

The ACR suggested that the real problem here might be that many performance measures ostensibly constructed from "samples" really are constructed from the complete set of actual observations. The ACR reasoned that frequently, a one-month observation is really a "sample" of the entire length of the production process, but is not a random sample, unless selected from among all of the months of production using some random procedure. In many instances, the proper statistical application may be statistical quality control viewing data as a time series. At the end of the six-month pilot test period, the confusion surrounding the sample versus population issue should be resolved. The ACR indicated that it would be very important to analyze the key underlying assumptions during the six-month pilot test period in order to establish the reasonableness of these assumptions and to understand the potential impact of any divergences from them.

Initially, the ACR plan did not contemplate a Z-test, or any other statistical test, for benchmark measures. It proposed to regard any measure that exceeds the benchmark value as a performance failure. Consequently, it envisioned that any performance worse than a benchmark would not be tolerated, and if exceeded, at least some penalty would be assessed. The ACR recommended monitoring the number of observations (e.g., orders) and improving benchmark measures over time taking into account the actual number of observations realistically expected to occur. For the immediate future, the ACR suggested treating benchmarks as absolutes, but moderating the impact of exceeding the benchmarks by means of smaller penalties for each occurrence. It also suggested that penalties should be greater for larger deviations from the benchmark.

Treating benchmarks as absolutes assumes that the parties established the benchmark values with some knowledge of the anticipated ability to meet them and / or the relative frequency of time they reasonably could be met. The frequency and value of the ILECs' inability to provide service meeting the benchmarks could be monitored and re-evaluated during the initial six-month pilot test period. Any dramatic differences between assessing performance with parity versus benchmark measures could eventually be resolved either by readjusting the alpha values, or benchmarks, or the incentives.

Minimum Sample Size

The ACR concurred with the concept of converting parity submeasures with rare activity to benchmarks. It suggested that additional rare activity submeasures should be converted to benchmarks. The ACR stated a preference for benchmark measures over parity measures for performance remedies, because benchmark measures do not require any complicating summary statistics. Early estimates indicated approximately forty percent of all measures were benchmarks, and that sixty percent were parity measures. Approximately fifteen percent of all measures had both parity and benchmark submeasures. The ACR expressed the hope that over time, the parties would agree to convert even more parity measures to benchmark measures.

The ACR surmised that sample size proposals were justified more by pragmatic concerns than by statistical principles. Proposed sample size specifications reflect negotiated values more than statistical criteria. For example, selecting a minimum sample size of five suggested one of two things: (1) either the cost to collect each observation is extremely expensive, or (2) there is an insufficient population from which to sample.[35] The issue of minimum sample size is relevant only for the first situation.

If all five observations occur during a particular time period, this is the entire population of measurements instead of a sample. The only sampling analog is to assume that the five observations are a sample of the potential observations that could have occurred during that same time period. Usually measurements are made with sufficient frequency to allow for corrective action if the process is beginning to "go out of control," or because management prefers to review data on a set periodic basis (i.e., hourly, daily, weekly, monthly, etc.). Such "periodicity" of measurement is usually established independent of sample size concerns. The ACR suggested that if too few observations occur in an established time interval, either the time interval can be lengthened, or the test can be performed using an aggregated measure incorporating more than one measurement. Or, the consistency of measurements could be tracked over time (e.g., number of "misses" for percent success measures) using statistical quality control charts.

The current assumption is that the time period for measurement is monthly. The ACR proposed lengthening the time period when the number of observations (e.g., sample size) is very small. However, the ACR recommended that this time period should not be so long as to enable the ILEC to manipulate results, and/or escape detection for providing non-parity service to the CLECs.

The ACR proposed to proceed with a minimum sample size of thirty, which could be aggregated in up to three-month time periods. Thus, a minimum

---

[35] For instance, such as might occur in the case where there are only five observations within a specified time period.

sample size of at least thirty would be obtained through an accumulation of up to three months, if necessary. If any sample size, aggregated or not, were to reach thirty in one, two, or three months, then the test would be performed when the number of observations first reached thirty. If, at the end of three months, the sample size had still not reached a minimum of thirty, the test would be performed using whatever sample size was achieved, regardless of the sample number. Ultimately, the measurement probably would be included in the rare occurrence benchmark list if fewer than thirty measurements happened during three months.

The ACR also advised that the appropriate length of time period for aggregation would be evaluated during the six-month pilot test period to better understand the frequency of measurements. Such an evaluation would aid in answering the question: "How many of each type of measurements can reasonably be expected to be made during any one month?" Any additional rare submeasures that could become benchmarks would also be evaluated during this pilot test period.[36] The ACR proposed to analyze any relatively large CLEC or ILEC values that skew the general tendency of the other values. (ACR at 24-25.)

Alpha Level/Critical Value

The ACR observed that it appears not to matter which critical value is actually employed, since the amount of the penalty can be adjusted to provide equivalent expected outcomes for the different possible critical values. The ACR proposed to track the actual alpha level outcomes, and ultimately calibrate the

[36] As stated, if there is no sample of observations, but instead, the population of CLEC values and / or ILEC values, the issues of errors and distributions are not really relevant.

size of payments as a function of the actual values. The greater the Z-statistic value (corresponds to a smaller Type I alpha error), the larger the penalty. The ACR proposed that in this proceeding, there should be no single critical cutoff value but a range of values. However, the ACR proposed that if one discrete cutoff value must be selected, it be a ten-percent Type I alpha level for parity tests. Preliminarily, ten-percent was a split between the suggested five and fifteen-percent values, and it is a commonly used critical value. This alpha level corresponds to 1.282 standard deviations.

The ACR described the CLECs' critical value proposal to be more of an "equal error" proposal than the "equal risk" proposal as the CLECs introduced it. Equal error refers to decisions with the same Type I and Type II error probabilities. Equal risk refers to decisions where the consequences of the decisions are equal, such as equal dollar losses. Their ultimate proposal does not equate the two expected dollar losses. In addition, the significance level that equates Type I and II errors varies by sample sizes and underlying distributions. The ACR also noted that the CLECs indicated concern with the Type II error, not just the Type I error. While fifteen-percent alpha levels are not commonly used for hypothesis testing, they are sometimes used for monitoring.

In their initial brief, the CLECs suggested that a performance payment be made for any occurrence beyond the acceptable level in a benchmark. (CLECs' 1999 Opening Brief at 3.) The ACR offered a similar recommendation, and pointed out that the CLECs also proposed that a specific table[37] be used to detail the small sample size benchmark standard comparable to the table agreed upon

[37] CLECs' 1999 Opening Brief at 33.

for large sample sizes (i.e., thirty or more observations). The ACR noted that the proposed table was negotiated, and did not systematically adopt the "closest" percentage possible compared to what would be expected from a large sample. It was unclear whether Pacific accepted this particular CLEC proposal.

The ACR remarked that while the concept of payments for all missed benchmark measures is easy to implement, it assumes accurate measurements. The ACR proposed discarding the benchmark table entirely at this juncture, and going with some level of graduated penalty for any measurement over the benchmark. For example, very small benchmark penalties could be assessed for very small frequencies of occurrences, and much larger penalties could be set for larger frequencies of occurrences.

For small sample sizes, the CLECs suggested permutation-testing procedures to compute the exact alpha and beta calculations.[38] (CLECs' 1999 Opening Brief at 30.) Pacific accepted this suggestion, specifying that the sample size should not be less than ten, if and when the Commission orders permutation testing. The company commented that permutation testing "is not an intuitive process for most people." Pacific recommended studying the validity and feasibility of utilizing permutation testing and that the approach be revisited

[38] Permutation testing involves direct estimation of probabilities from the actual data distribution, rather than inferences drawn from normal distribution "look-up tables."

after a trial test period. (Pacific 1999 Opening Brief at 2.) The ACR suggested that permutation-testing procedures might be a reasonable application.

Desiring larger numbers of observations so that there would be little need for permutation testing procedures as a result of sample size, the ACR outlined its concern. Proposed statistical procedures use one-tailed tests to indicate when penalties should be assessed against the ILEC for poorer service to the CLECs, but do not yield any incentives to the ILEC for providing exceptional service. Still, the ACR acknowledged that permutation-testing procedures could have some role in assessing more exact measures of error. The ACR recommended that during the pilot test period, there be an evaluation of this application of permutation testing.

The ACR asked the parties to respond to four specific questions[39] and to submit comments on the overall statistical model approach presented in the ruling. The parties[40] filed opening and reply comments on January 7, and January 27, 2000, respectively.

Responses to the ACR Questions and Comments on Its Overall Statistical Model Approach

Use of Standard Z-test or Modified Z-test

In response to the ACR's initial question of why the standard Z-test should not be used in the model, Pacific advocated retaining the Modified Z-test for three reasons. First, the standard Z-test yields inaccurate Type I error rates under the conditions apparent in performance remedies plans, i.e., in the absence

[39] The ACR questions are reproduced in the attached Appendix A.

[40] Pacific, the CLECs, Verizon CA and ORA.

of normal distribution and with relatively few large samples. Second, the Modified Z-test is easier to compute. Third, the Modified Z-test is sensitive to differences in the CLECs' variances. (Pacific Opening Comments on the ACR at 2 and 5.)

The CLECs urged using the Modified Z-test, yet agreed that the standard Z-test could be used. (CLECs' Opening Comments on the ACR at 3.) Verizon CA endorsed use of the standard Z-test, with modifications. It maintained that parties should be able to calculate and evaluate both the standard and Modified Z-tests during the evaluation or pilot test period. (Verizon CA Opening Comments on the ACR at 2.) ORA argued that since the underlying series or performance measures are not normally distributed, the true probabilities are unknown and the Z-test is of little value. It opposed using formal statistical tests for performance incentives. (ORA Opening Comments on the ACR at 5.)

Use of Benchmarks without Statistical Tests

To the ACR proposal to use benchmarks without statistical tests, Pacific asserted that benchmarks without statistical tests require an ILEC to meet an unreasonably higher standard of performance for small sample sizes than for large sample sizes. Pacific stated that statistical tests for benchmark measures make it possible to achieve a uniform Type I error rate for all measures under conditions of parity and compliance. Pacific segued from this question into an introduction of its white paper concept of converting all benchmarks to " standards" against which all the CLECs' results could be statistically tested. (Pacific ACR Opening Comments at 6 and 9.)

The CLECs remarked that the ACR's desire "to see more parity measures turned into benchmarks [ACR at 27] " was "troubling and difficult to understand." (CLECs' ACR Opening Comments at 8.) The CLECs continued to support a limited benchmark approach with no associated statistical component (except for the use of a table for small sample sizes). However, they maintained that unlike the parity standard, which requires the use of statistics to compare distributions, a benchmark standard requires no comparison other than the benchmark itself. The CLECs urged the enforcement of the benchmark standards adopted in D.99-08-020. (CLECs' ACR Reply Comments at 5.)

Verizon CA supported using benchmark measures without any statistical tests during the pilot period for all previously designated benchmark measures. Verizon CA agreed that the ACR's simple approach could be used during the pilot. Notwithstanding, Verizon CA proposed examining other alternatives such as tables for small sample sizes and the use of statistical tests with benchmarks. (Verizon CA ACR Opening Comments at 12.) ORA argued that benchmarks should be based on historical and not future data, and should be limited to those measures in which there is historical data available on at least 20 observations. (ORA ACR Opening Comments at 6.) ORA asserted that benchmarks should be defined as the historical mean of the series plus one standard deviation.

Use of Special Tables for Benchmark Measures

Pacific urged, and the CLECS agreed to, the use of special tables for percentage-based benchmarks with small samples. The CLECs favored the use of a table for benchmarks with small sizes. While allowing that the ACR's simple benchmark approach could be used, Verizon CA advocated alternatively examining the use of tables for small sample sizes. Verizon CA endorsed Pacific's adjusted table of percentages for benchmarks. As noted, ORA opposed the use of any formal statistical tests for performance measures.

Use of Minimum Sample Size of Thirty

In response to question 3, Pacific agreed that samples of thirty are adequate for average-based parity submeasures. It did not agree that a sample size of thirty is appropriate for benchmark measures that are interpreted as absolute standards and for percentage-based measures for which the benchmark is near zero (0) or 100 percent. (Pacific ACR Opening Comments at 12.) Pacific initially desired a minimum sample size of 30 occurrences, which is the standard "rule of thumb" for parametric statistical testing. As a compromise, Pacific was willing to lower the sample size to 20, with the caveat that the impact of the small sample sizes be evaluated at the end of the six-month trial test period. It also accepted benchmark measures for a specific list of rare submeasures, i.e., rare parity measures essentially become benchmark measures.

Pacific did not agree to use the sample size at whatever number of cases is available after three months if a CLEC does not have thirty cases. Stating that neither the CLECs nor the ILECs have proposed that sample sizes less than five (5) be considered for assessing remedies, Pacific did not want to set the minimum sample size at one (1) case. Pacific argued that aggregating over months introduces additional complexity and accounting expenses into the measurement reporting process and that a simpler rule for sample size examines results over one month. Pacific concluded that “while it may be possible to program these aggregation rules, they will make it difficult for the CLECs to monitor Pacific's performance and difficult for Pacific to manage its business.” (Id. at 13.)

The CLECs disagreed with using a minimum sample size as large as thirty (30). They argued that many CLECs would have fewer than 30 observations in a month for some measures. They also noted that Pacific reported that in the period of July through November 1999, approximately 100 CLECs had reportable data on 18,555 instances of parity submeasures. Of these reported submeasures, 62 percent of the CLECs had sample sizes of less than thirty cases. The CLECs further argued that the majority of all submeasures would have sample sizes less than thirty (30). (CLECs' ACR Opening Comments at 11.) Consequently, a majority of submeasures would not be subject to incentive payments. The CLECs have suggested a minimum sample size of 5 for parity submeasures. (CLECs' ACR Reply Comments at 8.)

The CLECs advocated using permutation testing for small sizes. (CLECs' ACR Opening Comments at 10.) They also disagreed with aggregating sample sizes over three months, or any time, because the ILECs could perform poorly for more than a month without correction. The CLECs insist that the only reason to favor a minimum sample size of thirty (30) for measured variables is that this might make a normal distribution an acceptable approximation to the distribution of the Z-test. Regarding minimum sample sizes for benchmark measures, the CLECs continued to advocate use of the table as the cleanest, easiest means of maintaining consistency with the adopted benchmarks. (Id. At 11.)

Verizon CA stated that aggregating small sample sizes over three months raises some potentially difficult and complex implementation issues. It advocated the standard Z-test with unequal variances employing exact distributions. For parity measures, Verizon CA favored using exact distributions for small sample sizes less than fifty (50). It also supported the Pacific-CLECs tables for benchmark measures with small sizes. Verizon CA disagreed that 30 observations for parity measures are appropriate with the Modified Z-test. It maintained that neither the standard nor Modified Z-test should be used with less than fifty observations. (Verizon CA ACR Opening Comments at 15.) ORA commented that the minimum sample size is not a "trivial issue" that should be arbitrarily set at thirty. It recommended a minimum sample size of 20 based on a formula (N (sample size) = 1 / a where a = .05). (ORA ACR Opening Comments at 8.)

Use of Ten Percent Alpha Level versus Fifteen or Five Percent

Pacific argued against the use of the 10-percent alpha limit and instead proposed a 5-percent Type I maximum error rate. The company asserted that a 10-percent alpha limit is unreasonably large and will yield an unfair proportion of Type I errors. It maintained that 5 percent represents a just compromise between unfairly detecting discrimination where none exists (Type I error) and failing to detect discrimination where it exists (Type II error). (Pacific ACR Opening Comments at 14-15.) Pacific focused on their desire to mitigate the effects of random variation. It commented that forgiveness rules help with the mitigation of random variation, but are complex and expensive to administer.

The CLECs continued to recommend an alpha value of 15 percent. They contended that it is a reasonable approximation of an alpha value that will balance Type I and Type II errors. The CLECs assert that they cannot ignore the impacts of a large Type II error. They also stated that any risk adjustment, such as a forgiveness plan, must reflect the alpha chosen by the Commission. The CLECs argued that an alpha value that more easily detects discriminatory behavior combined with a valid mitigation plan can achieve the goals of a high-powered test while minimizing payments under parity conditions. (CLECs' ACR Opening Comments at 13-14 and Reply at 10.) They agreed that there is no statistical reason why a 10-percent alpha cannot be used. In addition, they recommended that the Z-test for all parity submeasures be calculated throughout the six-month pilot test period at the five, ten, and fifteen percent levels in order to determine how many submeasures pass or fail depending on the critical value chosen. (CLECs' ACR Opening Comments at 13.)

Verizon CA commented that a 5 percent alpha remains a more balanced and reasonable choice. It asserted that a 10-percent critical value leads to a greater number of instances where a finding of "no parity" will follow from application of the test, when in fact, parity service is present. However, Verizon CA concurred with the CLECs that the result should be examined at all three proposed levels: five, ten, and fifteen percent. (Verizon CA ACR Opening Comments at 17 and Reply Comments at 4.)

ORA stated that an alpha level of 10 percent is simply too large. It argued that a more standard alpha level of 5 percent should be used. ORA stated that the use of a larger than normal alpha level means an increase in the probability of incorrectly declaring that the ILEC is out-of-parity. ORA urged the Commission to reject multiple alpha values as an attempt at data mining. (ORA ACR Opening Comments at 13.)

ORA also noted that the proposed remedies plan has no provision to prevent service deterioration, thus posing an unacceptable risk to ratepayers. It asserted that service levels can only be maintained if standards are based on prior historical data and not on future data. Performance measures used in the test period should be limited to those measures in which there is historical data available on at least twenty (20) observations. One of the two major goals that ORA identified for the Performance Remedies Plan is to maintain service levels at least at historical levels for all ratepayers. Its other goal is to ensure that customers of both the CLECs and the ILECs receive “statistically equal” service. Finally, ORA insisted that a benchmark should also be based on historical, and not present or future data. (ORA ACR Opening Comments at 6.)

March 2000 Workshop

In their reply, the CLECs recommended that the Commission hold a workshop on the new Pacific "white paper" proposal. ORA recommended that the Commission convene workshops to review all the various proposals. In all, the comments raised several issues requiring further discussion. To respond to the recommendations and address the unresolved issues, the assigned ALJ and staff facilitated a three-day workshop on March 27, March 28 and March 30, 2000. The workshop was divided into three daylong segments devoted to exploring the respective new Pacific and ORA plans, and further refining the components of a hybrid model.

The three segments of the workshop focused exclusively on the performance assessment part of the overall performance remedies plan (i.e., performance measurement, performance assessment, and incentive payment parts). The sessions did not include any substantive discussion of the performance measurement and incentive payment components of the remedies plan.[41]

For the purposes of the workshop sessions, the parties were to assume as given all prior work on performance measurements and benchmarks (on the separate parallel track pursuant to Commission Decision 99-08-020), including any current constraints. Parties were also to assume that any emergent performance measurement plan would use the performance measurements and

41 By ruling, the assigned ALJ advised the parties that the Commission would address the incentive components (including incentive structure, incentive amounts, and who receives incentive payments) after it determined the performance measurement and assessment plan components.

benchmarks resulting from the concurrent performance measurement phase of the proceeding. Finally, the parties were asked to delay incentive payment data modeling until the Commission selected the performance assessment model, or models.

The goal of the workshop was to develop more fully the three distinct performance measurement plans. These three plans were (1) the Pacific plan, (2) the ORA plan, and (3) a hybrid plan. All workshop participants were to assume on each specific plan’s day that they were advocates for that particular plan and that all participants would be jointly developing the "best" possible model for that specific plan type (i.e., hybrid, ORA, or Pacific). Where there were problems with various aspects of any plan, participants were asked to cooperatively recommend potential solutions for those deficiencies.

Participants also were asked to jointly determine if any of the plans were "fatally flawed" in any area, and if so, why. They were asked to follow the plan principles presented in the November 22, 1999 ACR, and to assume that the task before them would be to refine each particular plan type so as to be practical, capable of implementation and as simple as possible. Workshop participants were given an opportunity to advocate on behalf of their own plan on that specific plan's day, and to critique a competing plan on that plan's day. However, the intent of the sessions was to help refine each plan so that any one or all could be applied during the six-month pilot test period.

For each of the three plans, the assigned ALJ and staff focused on the respective model, element by element. There was a “straw man” or hypothetical proposal for each model element and either (1) a group decision was reached on that element or (2) a group modification was made to the hypothetical proposal. Discussion remained on each model subcomponent until a group "best" decision was reached, or it was evident that no decision could be reached and that the participants could only "agree to disagree." At the end of each plan subcomponent, a court reporter transcribed the group's findings on that plan element for the record.

Workshop Recommendations and Positions

Hybrid Performance Measurement Plan

At the workshop, Pacific, Verizon CA, the CLECs and ORA all agreed to use the Modified Z-test to develop a hybrid performance measurement plan.[42] Most of the parties also agreed that since they had selected the Modified Z-test, the use of a two-step standard Z-test procedure and other modifications[43] were no longer applicable in terms of the "Hybrid model." Verizon CA, however, supported using permutations, deltas and exact distributions in conjunction with the Modified Z-test.

The CLECs agreed to the initial hypothetical recommendation to treat benchmarks as limits without relying on statistical tests. Pacific and Verizon CA concurred with this as long as special tables based on statistical charts are used for all benchmarks. Pacific and the CLECs further agreed to produce two sets of consensus tables of acceptable misses for sample sizes scaled from 1 to 100 at a 10-percent alpha level. One set of the tables would represent percentage-based benchmarks, and the other would represent average-based

---

[42] Parties' consents to develop a hybrid plan did not imply their agreement with any resulting hybrid, as each party qualified its consent. For example, in response to the draft decision, ORA stressed that it did not support the hybrid plan or the Modified Z-test. ORA Comments at 6 (December 18, 2000).

[43] Such as the unequal variance Z-test, exact distributions, permutations, and deltas.

benchmarks. ORA opposed the proposition of treating benchmarks as absolutes without reliance on any statistical testing. (Reporters' Transcript (RT) at 1107, lines 10-24.)

All the parties assented to the second hypothetical Hybrid model recommendation that the Commission re-evaluate the benchmarks after a six-month pilot test period. However, Pacific's concurrence was subject to some preliminary data calibration occurring prior to the pilot. Moreover, the CLECs stressed that real penalties and incentives should be in effect during the six months.

In the discussion on sample sizes, Pacific, Verizon CA, the CLECs and ORA all supported the hypothetical recommendation that the time period for measurement of the sample be kept on a monthly basis. The second recommendation was that each party should precisely specify what minimum sample size it selects between five (5) and twenty (20). Pacific stated that it would go to a sample size of 5, with the proviso that there be mitigation measures to offset such a small sample size. Pacific further maintained that although it would apply the Modified Z-test for parity measures down to a minimum sample size of 5, it would not agree to use data or apply a permutation test below 5. Pacific argued that permutation testing was costly. In substantiation, Pacific agreed to submit operational cost calculations for permutation testing[44].

The CLECs selected a sample size of 5 and declared that if the minimum sample size were to be below 5, they would prefer permutation testing

---

[44] 2000 Pacific Workpaper #9

to be used. ORA would support a minimum sample size of 5; however, below 5 it would not support using the data. Verizon CA would support a minimum sample size of 20 with permutation testing. Below, Verizon CA would prefer to discard the data. Between 5 and 20 Verizon CA would prefer to use permutation testing, but without being subject to incentive payments. (Verizon CA May 5, 2000 Reply Br. at 5) Verizon CA strongly advocated permutation testing, and agreed to jointly submit with the CLECs after the workshop a description of a permutation testing protocol[45].

Following the 1999 performance incentive workshops, the parties identified six sub-measures[46] as "rare sub-measures." The parties purported to have agreed that there would not be an application of the minimal sample size to those measures or sub-measures identified as "rare." However, it was unclear from the briefs submitted after the 1999 workshops whether the parties still agreed as to what constituted the list of rare sub-measures. Thus, the third hypothetical sample size recommendation was to identify the measures or sub-measures upon which there was agreement that there would not be an application of the minimal sample size.

The parties agreed that rare measures or sub-measures would be those that rarely saw activity, yet were important to the CLECs. Pacific and the CLECs agreed to reanalyze the issue and submit as a workshop document any suggestions, additions or deletions to the group of six rare measures and

[45] 2000 GTE Workpaper #8.

[46] Sub-measure Nos. 26, 27, 30, 40, 41 and 43.

submeasures.[47] The rare measure list identifies those measures (or submeasures) where the measure would still be used at a sample size of one.

The parties also discussed how to make the Hybrid model operational for parity measures with no permutation testing and with sample sizes between five and twenty. To further the analysis, Pacific acceded to provide in two parts the "data on sample size for CLECs by submeasures." Pacific specified that one part of the analysis would show the percentage of the total data elements that would be used (not discarded). The second part of the analysis would show the percentages for the resulting sample sizes that would be used, relative to the entire set of samples. The company also offered to provide the absolute numbers, not just the percentages, for the previous two months of data.[48] (RT at 1135, lines 12-28.)

Pacific suggested that staff consider different remedy amounts when analyzing this data in the context of the "small sample world" versus the "large sample world." It questioned the reliability of the data if used with certain of the recommendations in the small sample realm. The CLECs proposed two recommendations to make the Hybrid model operational. First, small CLECs could be pooled into a sufficient aggregation to meet the minimum sample size. Second, a "mean plus standard deviation" similar to the ORA proposal could be used. (RT at 1136, lines 7-12.) Verizon CA supported the small CLECs pooling proposal, stating that it merited further exploration. (RT at 1136, lines 13-15.)

---

[47] 2000 Pacific/GTE Workpaper #10

[48] 2000 Pacific Workpaper #12.

Staff set forth two hypothetical recommendations on the Commission model's alpha level. Staff proposed that a 10-percent alpha level be used for the Modified Z-test. All the parties agreed to compromise at the 10-percent alpha level for the sake of developing the Hybrid plan. To the second proposal, that parties not calculate multiple alpha levels going forward, Pacific alone agreed to refrain.

In the January opening comments on the Hybrid model proposal, Pacific asserted that certain performance measures are based on failure rates for which no standard deviation has been defined. Thus, while a test similar to a Modified Z-test might be crafted for most of these measures, a Z-test could not be calculated for at least one of them as currently defined. (Pacific Opening ACR Comments at 5, footnote 5.) During the discussions on the Hybrid model the parties identified Measures 15, 16 and 19 as measures that might require special treatment or alternative application rules. At the conclusion of the Hybrid model discussion, Pacific, the CLECs and Verizon CA agreed to recommend a common solution to staff for these three measures.

Pacific, the CLECs and Verizon CA each detailed their respective lists of necessary enhancements to the Hybrid model. Pacific identified three necessary elements. The need to: (1) mitigate for random variations; (2) develop a procedure that deals with such excludable events, such as force majeure; and (3) establish an absolute cap for maximum exposure. Pacific noted that it was willing to pay up to $120 million in payments without evidentiary hearings in its latest incentive proposal. (Pacific ACR Reply Comments at Appendix 1.)

The CLECs maintained that their essential enhancement to the Hybrid model would be to convert all benchmarks associated with averages into percentage-based benchmarks. As a result, the benchmarks would be simplified and unified into one category.[49] Verizon CA specified five necessary enhancements to the Hybrid model. It would like the Hybrid model to either consider or perform correlation measures during the six-month trial period. Verizon CA would like the Hybrid to treat small sample sizes as they are being treated under the Bell South model.[50] It would also like the Hybrid model to consider real customer materiality[51] in contrast to statistical measurement differences. Verizon CA emphasized that all of the different measurement components are tied together, and some of its parts may have an aggregate effect that the Hybrid needs to consider. Finally, Verizon CA asked the staff to consider Pacific's white paper proposal as a tool to resolve many of the sample size issues or to satisfy the concerns about mitigation.

ORA Performance Measurement Plan

Foremost, ORA's plan attempted to adhere to the ACR's core guiding principal that any model under the Performance Remedies Plan be simple to implement and monitor. Thus, the first ORA proposal stressed simplicity as one advantage of its model. During the facilitated workgroup

[49] The CLECs stated that they would also be proposing this within the Performance Measurement Phase of this proceeding.

[50] "Statisticians for Bell South and AT&T have recently proposed a 'correction' to the Modified Z test that accounts for the skewness in the underlying distributions. They believe that this correction makes the 'modified-modified *t*' essentially equivalent to permutation testing at small sample sizes." Verizon CA Opening Brief at 23 (April 28, 2000).

[51] "Customer materiality" refers to whether the customer could actually perceive a difference between the performance to ILEC customers versus to CLEC customers, regardless of any statistical difference.

discussions, Pacific noted that while striving for simplicity was one of its concerns, there are more pressing substantive issues. The CLECs urged completeness and effectiveness in the remedies plan over mere simplicity. Verizon CA stated that the emphasis should be on fairness and accuracy, and simplicity should be one of several core principles. However, it asserted that if there were two plans equally effective and fair, it would prefer the simpler plan. Ultimately, Verizon CA suggested, ORA's plan may not be operationally simple.

ORA observed that the ILECs and the CLECs have proposed a mixed system with benchmark measures without any statistical tests to determine performance failure for some measures. ORA opposes using a mixed system. It argued that the same system should be applied to all performance measures, and that statistical tests are either relevant or they are not. (ORA ACR Opening Comments at 4.)

In its white paper proposal submitted in January, Pacific embraces the concept of a "same" system for both parity measures and benchmarks. However, Pacific asserts that benchmarks without statistical tests demand of the ILEC an unreasonably higher standard of performance (to avoid missing the benchmark) in the context of small sample sizes as opposed to large sample sizes. In contrast to ORA, Pacific asserts that statistical testing is relevant.

Both ORA and Pacific propose moving to a uniform system, but in different directions. The Pacific white paper plan advocates converting every performance measurement to a statistical test. The ORA plan advocates converting every performance measurement to a simple means and variance analysis, without any more formal statistical tests. The CLECs disagree that there is a need for a "same system." They contend that parity measures and benchmark measures need to be treated differently. Finally, Verizon CA notes that while the second ORA recommendation of consistency in terms of a "same system" concept is laudable, it is unnecessary.

The ORA plan argues that there are no provisions to prevent service deterioration in the Performance Remedies Plan. It states that current service levels can only be maintained if standards are based on historical, rather than future data. The current plans may have built-in reversed incentives such that if the ILECs were to increase the variability of their own processes, they could reduce incentive payments even though the CLECs receive worse performance. That is, the poorer the ILEC performs, the poorer the parity performance for the CLECs, but the larger variability would effectively prevent discrimination detection. To militate against this possibility, one of the straw man recommendations under the Hybrid plan was to monitor ILEC means and variances and compare them to historical values[52].

Responding to ORA's recommendation to base standards on historical data, Pacific questioned how the historical period would be defined and how the historical data concept could be operationalized. Pacific stated that it saw the ORA white paper as a conceptual approach that had not yet been specified to an operational level. It also requested a more detailed description of what monitoring ILEC means and variances would entail.

The CLECs advised that when one uses historical data in the context envisioned, there is a need for a lot of data. Overall, the CLECs were content with real-time data over historical data. However, they support monitoring the

[52] Pacific and Verizon CA agreed to provide staff with the incumbent local exchange carriers' historical means, variances and sample sizes for their retail parity measures and submeasures from September 1999 going forward through June 2000.

means and variances in order to mark improvement in Pacific's performance and to record where the CLECs stand in terms of Pacific's historical performance. Verizon CA noted that the data fluctuates substantially from month to month. Verizon CA maintained that there are inherent limitations in the depth and breadth of historical data necessitating adjustments. In addition, Verizon CA supported continuing to monitor the ILEC means and variances.

In its white paper proposal, ORA argued that neither Z-test, nor any other parametric test, should be used during the performance remedies plan six-month pilot period because many of the underlying performance measurement series are not distributed normally.[53] ORA argued that such abnormal distributions violate a fundamental assumption of the Z-test. Pacific supported using the Z-test during the pilot. It indicated a willingness to look again at the Z-test after the pilot, but wanted more specifics on what this would encompass.

Verizon CA commented that ORA's proposal to reject all statistical tests during the pilot is too extreme. Yet, it acknowledged that ORA's concern about normality was justified. Verizon CA suggested that ORA's approach should be considered after the six-month pilot is completed. At the workshop, Verizon CA cautioned that two factors should be taken into consideration. First, how to calculate the test statistics; and, second, how to use the calculation. Verizon CA noted that given assumptions of normality are met, one could consult "look-up" tables. Outside the range of normality, one could use permutation testing and exact distributions.

[53] As Pacific characterizes it: "no normal distributions and relatively few large samples." In fact, "samples" in question may not really be "samples," but rather time-series population observations.

The CLECs alone directly addressed the ACR's questioning of the use of any Z-tests. The CLECs recommend the use of existing parametric tests. However, they maintained that if actual experience does not justify confidence in the results, the test simply should be based on the number of observations that fall above some specified level. Essentially, this would convert measurement cases into counting cases. At that point, the CLECs propose to use the upper ten percent quantile[54] of the observed ILEC sample. CLEC statistical expert Dr. Colin Mallows of AT&T performed simulations and found that for some alternatives this non-parametric test is much more powerful than the Z-test. (CLECs Reply ACR Comments at 4-5.) In the workshop, the CLECs supported using "some flavor" of the Z-test during the pilot.

The ACR urged moving toward more aggregation of the measures over time in order to simplify the performance remedies plan. The aggregation effort should take all double counting out of the measures to the extent that there is correlation and interdependence between a number of the measures. In response, ORA stated that there are a total of 44 performance measures with over 1000 submeasures. It expressed concern about possible correlation between these measures. ORA argued that the ILECs' OSSs could be adequately measured with fewer performance measures, since many of the measures may be cross-correlated with each other and may not be needed. ORA's plan recommends that correlation tests be run for all the performance measures. It

54 A quantile is a portion of a distribution. An upper ten-percent quantile designates the highest ten-percent of results in a distribution, i.e., those results above the 90th percentile.

also submits that no performance measure should be included if it has a correlation greater than 0.80 with any other performance measure.

During the workshop, Pacific supported the hypothetical ORA plan recommendation for correlation testing. Pacific agreed that eliminating measures would help. To date, there has been no correlational statistical analysis or scientific modeling of the measures. However, given the contentiousness surrounding the issue, Pacific is willing to address the matter at a later time. Pacific admitted that the issues of correlation and interdependence had not yet been raised in the Performance Measurement Phase.

The CLECs pointed out that there was not a lot of data until recently to determine correlation. They do not want to get sidetracked with correlation issues at this point. While not adverse to a goal of reducing or adding measures if there is a legitimate rationale, the CLECs are opposed to a casual reduction of measures. They maintained that, at this point, Pacific and the CLECs see no further correlation between any of the submeasures. Verizon CA concurred with the plan recommendation as well as the ACR's desire to reduce the number of performance measures, if supported by the data. It asserted that the data is not currently available, and will not be available until after the six-month pilot. Verizon CA stated that the performance incentive phase would be the proper forum to address the issues of correlation and interdependence.

ORA's plan recommends a minimum sample size of twenty (20). It argues that a performance measure should only be used in the pilot if two requirements are met. First, that it satisfies a minimum sample size of twenty; and, that the measure is not highly correlated (greater than 0.80) with any other measure. At the workshop, Pacific, Verizon CA and the CLECs opposed ORA's recommended minimum sample size.

ORA's plan also recommends that parity be defined as a situation in which the average measured results for the CLECs served by a particular ILEC are within one standard deviation of the average measured results that the ILEC provides to its internal company units. ORA proposes that the ILEC average be based on historical and not future data.[55]

In terms of the workshop discussion, ORA's recommendation was to use the most recent historical fiscal or calendar year for the ILEC. None of the other parties supported ORA in its selection of one standard deviation. Assuming that ORA refers to one standard deviation of the mean, a one-tail test, and assuming normality, one standard deviation corresponds to approximately 84 percent the normal distribution, or a 16-percent alpha. However, this interpretation is somewhat inconsistent with the Office's prior recommendation of a 5-percent alpha, at least for large samples. For a one-tail test a 5-percent alpha corresponds to approximately 1.645 standard deviations. Assuming ORA was referring to the standard deviation of the mean, to facilitate the workshop discussions staff proposed using a 10-percent alpha or approximately 1.282 standard deviations for the sake of developing the ORA model.

However, a close read of ORA's proposal shows that ORA refers to one standard deviation of the observations, not one standard deviation of the mean.[56] In this case it is not possible to determine one critical alpha level

---

[55] In its comments on the draft decision, ORA states that historical data should be used for "the longest period for which data is available." ORA Comments at 7 (December 18, 2000).

[56] Statistical notation consistently used by ORA indicates its plan is based on one standard deviation of the observations: σ. *See* ORA Comments at 15 (December 18,

Footnote continued on next page

equivalent even with normal distributions, as one standard deviation of the mean is a function of the standard deviation of the observations and the sample size.[57]

ORA proposed that the benchmark be defined as the historical mean of the series plus one standard deviation. Consequently, any performance worse than the benchmark would trigger a penalty. ORA argued that the best demonstration of parity would be actual, not estimated, performance, even when experts using reasonable information make the estimates in good faith. The Office contended that proxies could be used in place of benchmarks in many cases. Since they are based on actual data, proxies are clearly superior to arbitrary benchmarks. ORA recommended that the Commission investigate their use before adopting arbitrary benchmark measures, and urged that benchmarks be used only in cases where there are no retail analogues and no proxies for those retail analogues.

Pacific, Verizon CA and the CLECs rejected ORA's benchmark proposal. They maintained that the reason they initially established benchmarks was because there were no retail analogues. Technically, there is no historical time-series data to calculate the mean and standard deviation for benchmarks under ORA's definition. Ideally, normalizing the benchmarks through proxies (assuming fairness and simplicity) is preferable to the current negotiated values. However, re-creating benchmarks distinct from those established in D.99-08-020

2000) and Reply Comments at App. A, at 4 (December 22, 2000). One standard deviation of the mean (standard error of the mean) would be designated: σ_m.

[57] *See* W. Hays, Statistics at 214-215 (5th ed. 1994).

would be impractical, contentious and time-consuming at this juncture. The parties accepted staff's recommendation to treat benchmarks as limits, as agreed to in D.99-08-020, in the context of the ORA plan.

Finally, staff asked the parties to help identify any other requirement conditions that need to be specified to make the measurement component of ORA's plan operational. In response, WorldCom introduced the "SiMPL Plan"[58] during the workshop. The SiMPL Plan would calculate the ILEC's historic performance percentiles and compare the relative CLEC performance results in those intervals. For example, non-parity would be identified if more than 10 percent of the CLEC's results were above the ILEC's 90th percentile. Other percentile comparisons would be made as well. WorldCom explained that this feature could assist in shaping CLECs' service expectations. It also contended that the plan is easy to administer since ILEC compliance is based upon whether the count of ILEC and CLEC events within each of three performance zones is proportional. (2000 MCIW Workpaper # 3 at 4.[59]) WorldCom characterized the SiMPL Plan as an alternative to the Modified Z-test in furtherance of the workshop assignment to collaboratively refine each model into the best that it could be. (Post-Workshop Opening Brief of AT&T, Covad, MGC Communications and WorldCom at 4-5.)

Pacific objected to WorldCom not presenting the SiMPL Plan in writing in advance of the workshop, and asserted that it saw "only minimal

---

[58] The Simplified Measurement of Performance and Liability Plan. 2000 MCIW Workpaper No. 3.

[59] Dr. George Ford's paper on the SiMPL Plan.

connections, at best" between the SiMPL[60] and ORA plans. (Pacific Opening Comments on Performance Remedies Workshop at 6.) Pacific described the SiMPL Plan as "fatally flawed"; simple only in that it does not require statistical testing to make the final determination of which measures were missed; and "inherently unfair to the ILEC." (Id. at 6-7.) Pacific concluded that the net result of the SiMPL Plan would be either to guarantee superior service to the CLECs or to plunge the ILEC into a spiraling series of costly service improvements that ultimately would not shield it from remedy payments. (Id. at 7.)

Pacific's White Paper Proposal

Pacific's revised Performance Remedies Plan, issued in January 2000, incorporates a number of new principles. First, Pacific maintains that there should be minimal payment of remedies when the ILEC provides parity service that is compliant with the standards of acceptable performance. This revised principal is similar to the ACR principal that "the plan should impose smaller penalties on Pacific for discriminatory performance that could be merely the result of random variation, and impose larger penalties for seriously deficient performance." (ACR at 12.) The ACR recognized this principal as a relative one, offset by benefits that the ILEC receives when it is not actually providing parity service but also is not measured as out of compliance.

Underscoring its first new principal, Pacific states as a supporting principal that the plan should not provide incentives for the ILEC to engage in behaviors to escape remedy payments other than performance improvements. It also insists that the plan should provide payment to the CLEC only for poor

[60] Pacific refers to it as "the Ford Model."

performance by the ILEC and not as a normal course of business. Further, Pacific restates the CLEC principal that the risks of Type I and Type II errors should be shared equally between the CLECs and the ILEC. Finally, Pacific asserts in its revised plan that samples of various sizes should be used provided the data they supply support valid decision rules.

Pacific's revised plan distinguishes between two definitions of parity service delivery. The company selected the definition that it contends recognizes and incorporates the variability of service delivery processes, i.e., the impossibility of delivering service exactly the same way every time. Thus, Pacific prefers the assertion that "parity of service delivery is achieved whenever the results for the CLEC and the ILEC are not *significantly* different." It notes that the key is to find a way to operationalize the meaning of "significant" when applied to ILEC and CLEC results. Pacific states that this is a statistical question that may be answered using models of the processes that produce the data to be evaluated. It is possible to calculate the probability of observing any particular difference between the results of the ILEC and CLEC given the assumption that parity service is being delivered. The probability of the observed difference in results is the mechanism for deciding the significance of the difference between ILEC and CLEC.

Pacific's white paper proposal advocates a definition of compliance that it maintains diminishes the disadvantages of measuring compliant service where there are no retail analogues. Instead of comparing CLEC results in absolute terms against a benchmark, CLEC results are compared in relative terms against a standard. CLEC results are compared to a standard using a statistical test to evaluate the compatibility of those results with the standard.

Consequently, "the results for the CLEC are compliant if they are not *significantly* different from the standard."

Pacific's revised plan contends that a key aspect of the use of standards and statistical tests is that the same criterion for the probability of failure (under the assumption of compliance) can be used as is used for parity measures. (Pacific's Opening Comments to the ACR, Attachment I at 4) Moreover, this probability can be made nearly constant for all sample sizes. Pacific disputed the CLEC's assertion that introducing standards at this late stage of the development of the remedy plan threatens to jeopardize all the difficult negotiations that went into the setting of benchmarks. The company insists that all standards may be derived from already agreed upon benchmarks by using a straightforward, objective formula.[61] The agreed upon benchmarks would remain intact and both sides would reap the benefits of using standards. (Id.)

In the revised plan, Pacific continues to propose a 5-percent alpha for parity measures. The white paper is not clear on what alpha level equivalent Pacific recommends for benchmarks with statistical tests. Pacific also contends that it is willing to go to a minimum sample size of 5 for parity measures, provided its white paper proposal for benchmarks is used. It recommends using the same minimum sample size of 5 for benchmarks.

Finally, Pacific recommends setting aside the forgiveness rules of its original plan, and sets forth an alternative mechanism for mitigating random variation. With this mechanism, Pacific proposes to focus on the CLEC as the unit of analysis and determine whether the total relationship between the ILEC

---

[61] Id. at 20, Appendix III.

and the CLEC shows evidence of discrimination or whether any failures observed can be ascribed to random variation. (Id. at 14.) Thus, Pacific would use a table to evaluate all the sub-measures for a single CLEC in lieu of forgiveness rules.

At the workshop, the CLECs disagreed with a performance assessment that uses statistical significance testing on benchmarks. They maintain that such a focus increases the complexity of the FCC's "a meaningful opportunity to compete" standard. The CLECs also contend that benchmarks are a surrogate for parity. Thus, benchmarks should not be treated the same way as parity measures. The CLECs support the existing treatment of benchmarks as tolerance limits not targets, as Pacific's plan would suggest. (RT at 1170-72.) Further, the CLECs continue to assert that there is a need for a mitigation plan for both Type I and Type II errors, and that all submeasures should be treated the same over time regarding both these categories of errors. (RT at 1170, lines 16-20.)

Verizon CA argued at the workshop that overall it supported Pacific's white paper model; however, it would like to see how certain specific elements of the model would be implemented. Verizon CA prefers permutation testing below a sample size of 50, and thinks the Modified Z-test down to a sample size of 5 presents problems. Within the context of the Pacific model, Verizon CA favors a 5-percent alpha and supports the concept of benchmarks with statistical testing. (RT at 1174, lines 7-24.)

ORA, noting concerns about the assumptions inherent in any parametric testing, reiterated that if the Commission adopts either the Hybrid or Pacific model we base them on historical data. ORA also suggested that we reassess the choice of alpha level, specific level of benchmarks, and the values of the small sample tables when more historical data becomes available. Moreover, ORA did not accept Pacific's argument that false negatives (Type II error) are unimportant because they do not harm the CLECs. It stated that performance incentives are fundamentally aimed at encouraging ILECs to provide parity of service and to dissuade attempts to discriminate, with the goal being to allow competition to proceed uninhibited. The fact that the attempted discrimination was unsuccessful does not mean that the performance incentive plan should not consider the attempt. (ORA Opening Brief at 3.)

Selection of the Decision Model

Our task now is to select a decision model consistent with several levels of policy goals. At the highest level, our model must effectively assist in converting a historical natural-monopoly market to a competitive market. This requires us to ensure that incumbents allow nondiscriminatory access to their infrastructures so competitors can provide local telephone services. That is, the CLEC's customers must not receive significantly worse performance from the ILEC than the ILEC's customers receive. Our decision today is at an even finer level of detail. We must specify a model that will accurately assess and identify discrimination. We must specify accurate calculations, accurate analyses, and accurate discrimination-identification decisions.[62]

We have reviewed the proposed models and the parties' comments regarding each of these models. While we had hoped that the parties would agree on a model and all the necessary implementation specifications, this did not occur. To the contrary, the parties disagreed on the models and on most of

---

62 We assume accurate data. Data accuracy is a topic in parallel proceedings.

their elements. While the workshop hybrid model[63] seemed to come closest to a successful compromise, the parties did not fully endorse it. At best, each party accepted the proposed hybrid model only insofar as we would modify it to address their particular interests.

Thus, we must review and approve or reject proposed models and / or elements, especially to resolve issues where there was no agreement. Unfortunately, virtually all model specifications by each party generated disagreement from at least one other party. The following is a list of the issues we must resolve now to specify the decision model for the next phase of this proceeding.

- Shall we select the workshop hybrid model, or any party's decision model, in its entirety, or should we select the best elements of different models to create a new hybrid?
- What statistical test[s], if any, shall be used to assess parity measures, including average, percentage, and rate measures?
- Where statistical tests are used, what decision criteria shall be used to identify results as parity or non-parity, or in other words, what criteria shall be used to identify test passes and failures?
- Shall a determination of material differences be a factor in non-parity identification?
- What sample size rules should be used?
- Shall data be transformed to closer approximate statistical test assumptions?

[63] When we refer to the "workshop hybrid model" we are referring to the outline model first described in the ACR, then subsequently revised in the workshops. Beginning with modifications in the workshops, this model was referred to as the "hybrid model" since it incorporated components from the different models.

- Shall benchmarks be used as limits or as targets, and shall statistical tests, or tables based on statistical analyses, if any, be used for: (1) Some benchmark measures, (2) All benchmark measures?
- Shall correlational analyses be employed to assess and reduce redundancy between performance measures?
- Shall historical data be used as a decision criterion, or be monitored separate from the identification of passes and failures?
- Shall existing benchmarks be modified to address new developments in this assessment phase of the proceeding?
- Should we specify different models for the different ILECs?
- Should we plan to adjust payments retroactively after the six-month trial period?
- What other specifications should we order to enhance the use and understanding of our decision model?

We will base this decision on the following criteria:

- Accuracy: Identify discrimination when it exists, and do not identify discrimination when it does not exist.
- Correctability: When more important criteria do not provide conclusive guides to our decisions, we will select the elements that offer the most opportunity for correction in later phases of this proceeding.
- Academic soundness: Our rationale shall be based on recognized applicable statistical assumptions and principles, and confirmed by data when possible.
- Policy goals: Our rationale will be consistent with competition-enhancing policy and law providing substantially equal access for all potential local phone service providers, whether small or large.
- Simplicity: Without sacrificing higher-order goals such as accuracy, we will prefer the more simple models and elements.
- Fairness: We will strive to be as even-handed as possible to optimize competitive market potential and benefits.
- Openness: We will document and explain the criteria we use in selecting the model and its elements so that all parties can

knowledgeably comment and knowledgeably argue for modifications to the model.

- Consensus: We will prefer models or elements where a consensus exists, unless there are differences on more important criteria.
- Experimentation: Rather than consider the initial model to be a final product, we will consider this initial implementation to be an experiment that will inform future model development.
- Costs: Unless a more costly model or element is likely to better satisfy important criteria, we will prefer less costly approaches.
- Understandability: When differences on more important criteria are minimal, we will prefer more easily understandable models and elements. We will also take care to explain models, elements, and analyses in sufficient detail and at a level to help the reader understand the model we specify and the reasons we have selected the model and its elements.

From the parties' proposals and comments, relevant statistical sources, and staff's analyses, using the above criteria we have selected a decision model.[64] The model is presented in Appendix C. The following is a discussion of the model and our rationale for selection of the various model elements.

Decision accuracy

While the above criteria lists may seem self-explanatory, we believe it important to discuss at length the first and most important criterion, decision accuracy. We begin with a brief overview.

Once performance measures are established and results are obtained, accurately assessing the existence of competitive conditions then becomes a

[64] Accordingly, we take official notice of several academic sources. They are referred to throughout the following discussion and are listed in Appendix B. Additionally, we take official notice of several analyses performed by staff which are included as appendixes to this Decision.

decision-making task. Since these decisions must be self-executing, the Commission must construct a decision model that can automatically identify performance result levels that reveal competition barriers and that will trigger incentive payments. There are two fundamental categories of performance measures that must be assessed. These categories' definitions are based on the characteristics of the service an ILEC provides a CLEC and the CLEC's customers. Where there is an ILEC retail analogue to the service given the CLECs and their customers, the FCC has stated that parity of services is evidence of open competition.[65] Where there is no ILEC retail analogue to service given the CLECs, then open competition is gauged by performance levels that provide a "meaningful opportunity to compete."[66] These performance levels that have no retail analogue are designated "benchmarks." Thus, the two categories of measures have been termed "parity" and "benchmark" measures.

Decisions regarding parity measures

In identifying parity or non-parity, accurate remedies-plan decision-making is not simply a matter of accurately calculating average ILEC and CLEC performance and identifying non-parity if ILEC service to CLEC customers is worse than ILEC service to ILEC customers. Given that there is variability in ILEC performance in its own retail services to its own customers, a measurement

---

[65] Parity of services refers to "access to competing carriers in 'substantially the same time and manner' as it provides to itself" and "access that is equal to (*i.e.*, substantially the same as) the level of access that the BOC provides itself, its customers, or its affiliates, in terms of quality, accuracy, and timeliness." *Bell Atlantic New York Order* ("FCC BANY Order"), 15 FCC Rcd at 3971, ¶ 44.

[66] *Id*. at 3971-72, ¶ 44-45.

result of inferior service to CLEC customers could be due either to this variability, or actual discrimination, or both. In other words, if we sample the ILEC's service results to its own customers, we will get different results, some better, some worse than the average. Service to a CLEC may be viewed as a "sample" of the ILEC's services.[67] Theoretically speaking, if the performance measured from the CLEC "sample" is typical of the performance for similar ILEC customer "samples," then there is no evidence of discriminatory service, even if it is somewhat worse than the ILEC average. However, if the CLEC "sample" performance is worse than most ILEC customer "samples," then there appears to be evidence of discrimination.

In statistical terminology, the non-discriminatory variability between multiple ILEC samples is termed "sampling error" or "unsystematic variability," referring to the fact that the variability is simply due to random sampling outcomes. Discriminatory variability is the case where the performance in a CLEC sample is worse than what would be reasonably expected from sampling error. Discriminatory variability is variability that goes beyond sampling error and is termed "systematic variability," meaning that something is systematically causing the differences between the samples. Since these two types of variability cannot be directly observed, discrimination or non-discrimination must be indirectly inferred.

---

[67] By using the word "sample" we do not mean to imply that the correct model is a sampling model in the traditional parametric statistical use of the term. The record does not help us resolve, nor do we resolve, the underlying assumptions of whether the combined ILEC and CLEC performance results for each month should be viewed as a "sample" of an underlying process distribution, or if each month's results should be viewed as the entire population of events.

A decision outcome matrix illustrates this problem. Figure 1 presents the four possible decision outcomes about parity. The four outcomes represent conclusions of either parity or non-parity of service under conditions of either actual parity or non-parity. The decision outcome matrix simply recognizes that when we make a dichotomous decision, there are four possible outcomes, two correct and two incorrect. In the context of this proceeding, the decision outcome matrix illustrates decision goals: (1) to detect differences when they exist, and (2) to not detect differences when they don't exist.

Figure 1: Decision Matrix

	Parity Identified (Decision: No Discrimination)	**Non-Parity Identified** (Decision: Discrimination)
Reality: Parity (No Discrimination)	**Correct Decision** (True Negative)	**Incorrect Decision** (False Positive)
Reality: Non-Parity (Discrimination)	**Incorrect Decision** (False Negative)	**Correct Decision** (True Positive)

Figure 2 expands this illustration. Given that decisions regarding parity are based on measurements that are comprised of both "true" values and "error," these outcomes can represent both correct and incorrect decisions, depending on the relative amount of error in the measurement. Figure 2 portrays sampling error effects.

Figure 2: Decision Matrix Showing Sampling Error Effects

	Decision: Parity	Decision: Non-parity
Reality: Parity (No discrimination)	**Correct Decision** Relatively low sampling error	**Incorrect Decision** Sampling error creates spurious difference
Reality: Non-parity (Discrimination)	**Incorrect Decision** Sampling error masks real difference	**Correct Decision** Relatively low sampling error

Figure 3 illustrates the contribution of statistical testing. The potential for errors is the same as in the first two matrices where no statistical testing is applied. The only contribution of statistical testing is that it allows us to estimate decision accuracy, or in other words, to calculate the decision error probabilities. These probabilities can then assist decision-making by quantifying the different error probabilities and comparing them to standards of confidence that we wish to apply. These standards of confidence are expressed as: (1) the power of the test, and (2) the confidence level.

Figure 3: Decision Matrix with Statistical Tests

	Decision: Parity	**Decision: Non-parity**
Reality: Parity (No discrimination)	**Confidence level** Probability = ***1 – alpha***	**Level of significance** Probability = ***alpha*** Type I error
Reality: Non-parity (Discrimination)	**Test insensitivity** Probability = ***beta*** Type II error	**Test power** or sensitivity Probability = ***1 – beta***

Test power refers to the ability of the test to actually find true differences, that is, the confidence that you found what you were looking for, when it existed. "Confidence level"[68] refers to the ability of the test to reject spurious differences, that is, the confidence that when you identified something, it actually existed. Together, these probabilities represent the amount of confidence one can have in decision quality. The higher the test power, the greater the confidence one can have that true differences were uncovered. The higher the "confidence level" the greater confidence one can have that discovered differences are real differences. Other things being equal, as one level of confidence is increased, the other decreases. In other words, the more powerful the test, the more likely there will

---

[68] While by convention 1–alpha has been termed the "confidence level," in reality both 1–alpha and 1–beta are confidence levels. They are distinguished by the type of confidence they estimate.

also be differences found solely due to random variation, and the higher the confidence level, the more likely true differences will be missed. Neither confidence standard is inherently more important than the other. Each application of a statistical test implies different trade-offs between these two confidence standards, and their corresponding error probabilities, depending on the consequences of the two different errors.[69]

In the present case of restructuring a historical natural-monopoly market to create a competitive market, the primary function of performance measurements and the decisions about performance measurements is to detect and prevent barriers to competition. To maximize goal attainment these decisions must be as accurate as possible, to find and prevent actual barriers, and to avoid identifying barriers when they do not exist. However, there is no legislative or regulatory guidance specifying the relative importance of the two decision errors.

On one hand, if we do not detect barriers when they occur, competition may fail, and the fundamental purpose of the legislation will have been thwarted. On the other hand, if we identify barriers when they do not exist, then we are likely to take unfair punitive action. Therefore we will use statistical testing to assess the balance between these two competing outcomes, thus enabling greater decision quality and attainment of legislative goals. Figure 4 summarizes the statistical decision matrix and identifies the probabilities that correspond to the four possible decision outcomes.

---

69 *See* W. Hays, Statistics at 267-303 (5th ed. 1994), and B.J. Winer, Statistical principles in experimental design at 10-14 (1971). We discuss these issues in more detail in a following section.

Figure 4: Decision Matrix Statistical Testing Summary

	Decision: Parity	**Decision: Non-parity**
Reality: Parity	No barriers exist. No barriers identified. (***1 – alpha***)	No barriers exist. Barriers identified. (***alpha***) Type I error
Reality: Non-parity	Barriers exist. No barriers identified. (***beta***) Type II error	Barriers exist. Barriers identified. (***1 – beta***)

Using measures of performance averages and variability, statistical analysis provides estimates of: (1) the probability that a result of a certain magnitude would be detected when it exists (test power and corresponding error *beta*) and (2) the probability that the result is due to random variation when in fact there are no differences (confidence level and corresponding error *alpha*). The methodology for using these estimates to establish dichotomous decision criteria is called null hypothesis significance testing. The analyst specifies a null hypothesis to pose that there are no differences between two performance outcomes, selects a confidence level that strikes the appropriate balance between the two types of error, calculates the probabilities, and compares them to the selected significance level. If the probability is less than the selected significance level, then the analyst rejects the null hypothesis and accepts the alternative hypothesis that there are real differences.

In the two approved Section 271 applications to date, Bell Atlantic New York and Southwestern Bell in Texas use a "Z-test" statistic to calculate these probabilities. Conceptually, the Z-test statistic compares the ILEC's average (mean) performance to the CLEC's mean performance, and then compares the difference between the means to the difference that would be expected from random variation at a selected confidence level. The expected difference is calculated from the variability in the samples of performance. The greater the variability, the greater the expected difference, and the less likely a true difference will be detected. In the Z-test, the difference between means is compared to (actually divided by) an expected difference term that is calculated from the sample size (n) and the variability in those samples (variance).

Thus the sample size, the variability in the samples, the power of the test, the confidence level, and the size of the true differences between means affect decision quality.[70] These elements are interdependent such that changing one will have an unavoidable effect on at least one of the others. A convention has existed for several decades to pre-select a fixed confidence level (or alpha) and adjust the other elements if desired. For example, if a test with the common 95% confidence level (0.05 alpha) lacked adequate power to detect true differences, the sample size could be increased. Methods have been developed to calculate the minimum sample size required to attain adequate test power.[71]

Additionally, since much of science depends on replication, test power is relegated less attention because of the expectation that experiment replication will address this issue. However, this convention which evolved in the 1920's, called null hypothesis significance testing, has been questioned over the last

---

[70] W. Hays, *supra* at 289-293 (1994).

[71] For example, *see* W. Hays, *supra* at 333-334 (1994).

three or four decades. At least one professional standards board was recently established to consider abandoning such testing in favor of new methods that strike a more even balance between test power and confidence levels.[72] Illustrating this concern about ignorance of test power, the following comments reveal some of the intense dissatisfaction with current research relying on 0.05 critical alpha levels:

> Whereas most researchers falsely believe that the significance test has an error rate of 5%, empirical studies show the average error rate across psychology is 60%--12 times higher than researchers think it to be. The error rate for inference using the significance test is greater than the error rate using a coin toss to replace the empirical study. . . . If 60% of studies falsely interpret their primary results, then reviewers who base their reviews on the interpreted study "findings" will have a 100% error rate in concluding that there is conflict between study results. (p. 3.)[73]

The balance between these interdependent elements that affect decision outcome quality is problematic not only in pure research contexts, but also in applied contexts such as engineering and operations management.[74] As parties have greater vested interests in different outcomes, the greater the argument

---

[72] R. Hubbard; R. Parsa; M. Luthy, The spread of statistical significance testing in psychology: The case of the Journal of Applied Psychology, 1917-1994, 7 Theory & Psychology at 545-554 (1997).

[73] J. Hunter, *Needed: A ban on the significance test*, 8 Psychological Science at 3-7 (1997).

[74] For example, *see* C. Das, *Decision making by classical test procedures using an optimal level of significance*, 73 European Journal of Operational Research at 76-84 (1994); R. Verma & J. Goodale, *Statistical power in operations management research,* 13 Journal of Operations Management at 139-152 (1995); and K. Brubaker & R. McCuen, *Level of significance selection in engineering analysis*, 116 Journal of Professional Issues in Engineering at 375-387 (1990).

there is over the appropriate balance. This is certainly the case in the present proceeding. Parties disagree on what is appropriate for all elements: the appropriate tests, confidence level, test power, sample size, test statistic, and other elements and nuances of a statistically based decision structure.

Determinations regarding benchmarks

Unlike performance measures where there is a retail analogue, benchmarks cannot compare ILEC service to CLEC service since there is no ILEC service analog. Instead, benchmarks are judgments about the levels of ILEC performance for CLEC competitive service that are necessary to "allow a meaningful opportunity to compete." Benchmarks have been constructed as tolerance limits. For example, one measure specifies that *99 percent of billing invoices shall be available within 10 days of the close of the billing cycle*.[75] The issues for statistical analysis accuracy are not the same as for parity measures. However, small sample benchmark applications raise similar decision matrix issues that we discuss after we address the more complex issues of the statistical models for parity performance measurement results.

Statistical models

As discussed, several models for parity assessment have been presented during the course of this proceeding. Some were intended to be complete, such as Pacific's most recent model. Other models were intended to present conceptual frameworks that would resolve various problems and which could be implemented with further negotiation and development. Examples of these

---

[75] Performance measurement No. 30, Wholesale Billing Timeliness, D.99-08-020, *mimeo.* at 43.

include ORA's model, MCI's SiMPL model, and the ACR's proposal.[76] We find that none of the presented models are acceptable in their entirety. Our rationale for this finding is best explained by discussing our evaluation and selection of the model elements that we will specify in what will be a new "hybrid" of elements from each of the different models presented in this proceeding.

Statistical tests

Three types of parity measurements have been developed for monitoring ILEC performance: averages, percentages, and rates. Each measurement type requires a different statistical test or a variant of the same test.

Average-based measures

The choice of a statistical test for average-based parity measures came as close as any model element to being accepted by all parties. Pacific and the CLECs have agreed that the Modified Z-test should be applied to average-based measures. Verizon CA also agreed to use the Modified Z-test, albeit with modifications. Only ORA disagreed, although they consented to its use in the development of a "hybrid" model. (RT at 1103.) All parties have agreed that a one-tailed test should be used. A one-tailed test is appropriate for situations where we are only interested in outcomes in one direction, in this case where the CLEC performance results are worse than the ILEC results. This is consistent

[76] In comments to the draft decision ORA asserts that its proposal specifies an implementable model. We appreciate ORA's sincere efforts to present a simplified model which is intended to avoid recognized problems with other models, such as data distribution non-normality. However, ORA's proposal leaves unclear critical components, such as calculation of the "standard deviation" as discussed *supra*. If ORA wishes to explore its proposal further, we urge them to present explicit formulas and data examples to the other parties, and ultimately, to us during the next phase.

with academic texts[77] and with the FCC's view of the appropriate statistical application regarding the requirements of the Act.[78]

Standard Z-test

The standard Z-test compares the difference between means to what is essentially an expected difference between means that could be explained by random variation. The expected difference is calculated from the variation (variance) in both the ILEC and CLEC results. The ACR proposed that the ILEC and CLEC variances be screened for statistically significant differences as a first step, then either the pooled or equal variance standard Z-test statistic would be calculated as a second step depending upon the results of the first step. Verizon CA described several concerns with the ACR's proposed two-step standard Z-test method and suggested several corrections.[79] However, in response to the CLECs' concerns that ILEC discrimination could increase the CLEC variance, and thus make it more difficult to detect any discrimination, all parties agreed to use a Modified Z-test instead of the standard Z-test.

Modified Z-test

This test was first adopted by the NYPSC for the BANY 271-application performance remedy plan.[80] Similar to our situation, since the CLECs were concerned that by providing highly variable service to the CLECs, the ILEC

---

[77] Hays *supra* at 293-294 (1994); and Winer *supra* at 20 (1971).

[78] *Bell Atlantic New York Order*, 15 FCC Rcd at 4191, App. B, ¶ 18.

[79] Verizon CA ACR Opening Comments at Apps. A and B (January 7, 2000).

[80] *Bell Atlantic New York Order*, 15 FCC Rcd at 4182-4188, App. B., ¶¶ 1-13.

theoretically could increase the expected difference and thus mask real differences, the parties in the BANY application proceedings agreed that the CLEC variance would not be part of the expected difference calculation. This alteration has been given the name "Modified Z-test." The FCC considers this test reasonable,[81] and it has been favorably presented in statistical academic literature.[82] The FCC subsequently approved Southwest Bell's performance remedy plan for Texas, which also uses the Modified Z-test.[83]

Only ORA objects to use of the Modified Z-test, although for the purposes of developing a hybrid model, ORA is willing to proceed using the test. (RT at 1103.) ORA's primary concern is based in their opinion that use of any Z-test requires that the data be normally distributed. According to the statistical literature, this may be only partially correct; Central Limit Theorem states that for sufficiently large samples, non-normality in the data does not affect the test.[84] With large samples, the distribution of sample means will be normal, whether or not the raw data distribution is normal. The means of

---

[81] *Id.* at 4188, App. B ¶ 13 and n. 37.

[82] C. Brownie, D. Boos & J. Hughes-Oliver, *Modifying the t and ANOVA F tests when treatment is expected to increase variability relative to controls,* 46 Biometrics at 259-266 (1990).

[83] *See* SWBT interconnection agreement, *Texas T2A Agreement*, Attachment 17: Performance Remedies Plan, ¶ 2.0 at 1.

[84] "If a population has a finite variance σ^2 and a finite mean μ, the distribution of sample means from samples of N independent observations approaches the form of a normal distribution with variance σ^2 / [sqrt(N)] and mean μ as the sample size increases. When N is very large, the sampling distribution is approximately normal. Hays (1994) at 251. See also, R. Khazanie, Statistics in a world of applications at 344-345 (4th ed. 1997).

sample sizes of 30 or more are typically considered sufficiently normally distributed to have minimal effect on a Z-test.[85] The BANY performance remedy plan addresses this issue by using the Modified Z-test down to a sample size of 30, and is temporarily using the t-test for smaller samples until permutation testing is established.[86] In comments to the draft decision, ORA asserts that only its proposal is consistent with Central Limit Theorem. ORA Comments at 9. We are not persuaded and remain concerned that no proposal has adequately addressed what a "sufficiently large" sample is. For example, ORA states that over time, distributions will approach normality because the number of observations will increase. However, there is no evidence that the distribution of the *observations* will be normal for very large samples.[87] Our understanding is that only the distribution of sample means will approach normality as sample sizes increase. Yet even ORA's model appears to depend on results limited to a month interval. ORA Opening Comments on the ACR at 9. Additionally, the adverse affects of non-normal data may be quite limited. For example, a statistical text cited by ORA to support its views on Central Limit Theorem also states,

---

[85] *Id.* at 349-351.

[86] *Bell Atlantic New York Order*, 15 FCC Rcd at 4187, App. B., ¶ 11. We assume that the t-test used by BANY is the Modified Z-test with the resulting Z-statistics compared to critical values in a t-distribution table rather than a normal curve table. *See also* Khazanie, *supra*, at 410-411 (1997), and Brownie, et al., *supra*, at 260-261 (1990).

[87] See the graphs and data tables presented in conjunction with the discussion herein of data transformations. The presented data is actual commercial performance data. It is extremely non-normal even at sample sizes of as large as 179,000 cases.

> Regardless of the shape of the population from which we draw our samples, the sampling distribution of means will be normal *if the sample size is sufficiently large*. What is a "sufficiently large" sample? There is no easy answer, because the required sample size depends on the shape of the population distribution. You will find some statistics texts specifying an *N* of 30 and others an *N* of *50*; certainly an *N* of 100 would remove all doubt about the resultant shape of the sampling distribution. In any event, the central limit theorem enables us to solve problems without worrying whether or not the population from which we are sampling is normal. (p. 151, italics in original text, underlining added.)[88]

We appreciate ORA's persistence in raising this concern, and agree insofar as we acknowledge that non-normality is a problem of an unknown extent. We will not act on this until we receive more evidence on the extent of the problem before prescribing for the final decision model any statistical tests that may be adversely and meaningfully affected by non-normality.

Verizon CA agrees to use the Modified Z-test, although its agreement is conditional. Most importantly, Verizon CA agrees to use the Modified Z-test for average-based measures if a permutation test is used for small samples. As discussed below, we agree with the concept, but have concerns with the implementation.

Permutation tests

To remedy the problem of small samples, which may not meet the "normality" assumptions of the Modified Z-test, Verizon CA proposed that a permutation test be used for average-based and other performance measures. The permutation test is a statistical test that, independent of any underlying

[88] A. Bartz, Basic Statistical Concepts at 150-151 (1988).

distribution, assesses the probability of an outcome. As such it is termed a "distribution free" or non-parametric test in contrast to the parametric Z-test which is based on distribution assumptions.[89] The reasoning behind its use is that when the Z-test normality assumption is violated, a permutation test is more appropriate and accurate since it compares the actual CLEC data directly to the ILEC data without making distribution inferences. Theoretically, the test is only necessary for smaller samples where Central Limit Theorem does not predict normality, because the two tests should produce similar results for larger samples. Differences in distributions do not affect permutation test results, and "look-up" distribution tables, such as "Z" or "t" tables are not necessary.[90] In theory, the benefit of permutation testing is that it can increase the accuracy of the error estimates, thus enabling more accurate decisions.

Only Pacific objects to the use of permutation tests.[91] Pacific originally objected to the assumed costs of such a procedure, but continues to object even though those costs have turned out to be much smaller than originally assumed.[92] Pacific now objects to the procedure as being inadequately

---

[89] *See* generally, P. Good, Permutation tests: A practical guide to resampling methods for testing hypotheses (2nd Ed. 2000).

[90] *See* Mallows Aff., FCC CC Docket No. 98-56, ¶¶ 25-29 at 15-17 (May 29, 1998).

[91] Pacific Reply Brief at 14-15 (May 5, 2000).

[92] Pacific originally estimated the implementation cost of permutation at .75 to 1.2 million dollars (Pacific Bell response to staff questions, February 11, 1999 workshop). Recently Pacific updated their estimate, showing a $300,000 initial implementation cost, with $24,000 to $36,000 yearly maintenance and operational costs (Pacific Bell, deliverable no. 8, April 13, 2000), although we are not aware of any competitive bids that might serve to reduce this estimate further.

tested and too complex,[93] although earlier had acknowledged its feasibility at least for Pacific samples less than 5,000 or 10,000.[94] Regarding the feasibility of its use for such large samples, Verizon CA has presented procedures for implementing permutation on samples of any size.[95]

The selection of the appropriate test for small samples should be based on the relative accuracy of the different tests. The permutation test has the potential for being a more accurate test that can handle small samples. Contrarily, the Z-test relies on the resulting sampling distributions being normal. Evidence in this proceeding is compelling that normality cannot be assumed for small samples since measures of time-delay are commonly skewed – the distribution is "bunched up" for shorter delays, and tapers off slowly for longer delays. (*See* Figure 5 for a hypothetical example of a provisioning frequency distribution.)

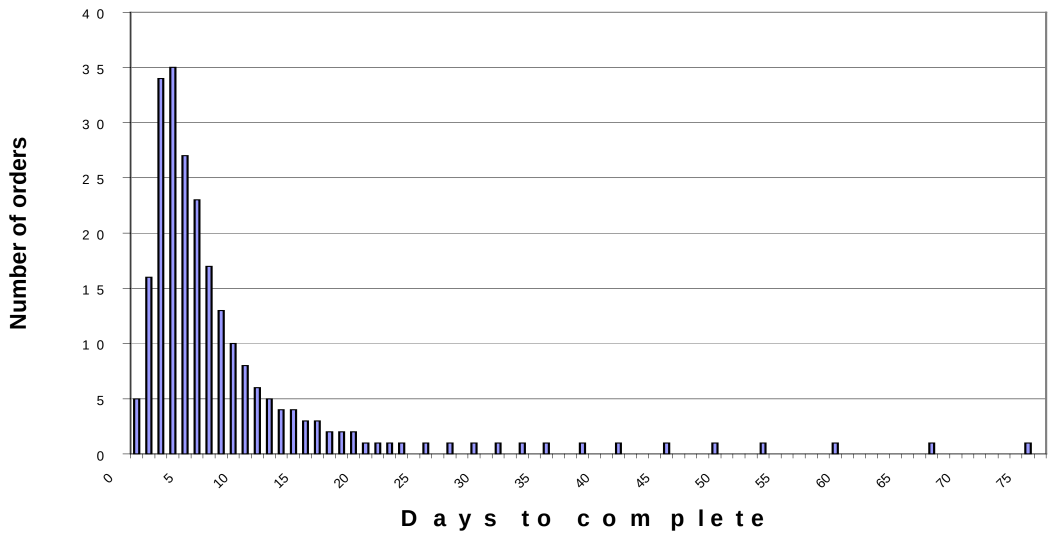
---

[93] Pacific Reply Brief at 14-15 (May 5, 2000).

[94] 2000 Pacific Workpaper No. 9 (April 13, 2000).

[95] I.e., resampling techniques. Verizon CA Opening Brief, Attachment 1 at 1 (April 28, 2000). *See also Bell Atlantic New York Order*, 15 FCC Rcd at 4189, App. B, n. 38, and P. Good, *supra* (2000).

Figure 5: A skewed distribution



Given the Z-test's problems with non-normal data, and the fact that the permutation test is unaffected by different distributions, it is possible that the permutation test will be more accurate, and thus would be the preferred test. Theoretically, one should expect that the permutation test would calculate alphas that diverge from Z-test-produced alphas increasingly as sample sizes decrease – the smaller the sample, the larger the discrepancy. On the other hand, as sample sizes increase, the alphas from the two methods should converge toward equality for large samples. Unfortunately, the few data examples we have available to us do not show this expected relationship.[96] The examples show the expected divergence for small samples, but not the expected convergence for larger samples, contrary to the theoretical expectation that the

[96] John D. Jackson, *Using permutation tests to evaluate the significance of CLEC vs. ILEC service quality differentials*, Verizon CA Opening Brief, Attachment 1 at Appendix 2 (April 28, 2000).

results should be the same for large sample sizes.[97] These results raise doubts that the record before us is sufficiently developed to allow us to confidently select the permutation test as a superior test. Either the permutation test is treating data differently than we would expect, or a sample size of 30, or even 131, is still too small to expect sample mean distribution normality for these performance measures. We note that the permutation test is relatively insensitive to outliers[98] compared to the Z-test. This insensitivity occurs because in the final step, the permutation test treats the data as ranked data where an extreme score's value does not influence the outcome.[99] In contrast, extreme scores influence the Modified Z-test.[100]

This result raises the question whether extreme scores would have insufficient influence in a permutation parity test, insofar as these extreme scores might be some of the most publicly noticeable indicators of

---

[97] *See* Jackson, *supra*, at 2-9.

[98] In this application a statistical outlier refers to rare extreme scores, for example, a large but rare performance failure such as an unusually long provisioning time.

[99] R. Khazanie, *supra*, at 720 (1997).

[100] This insensitivity can be illustrated by examining the data example originally presented by Dr. Mallows, but elaborated by Dr. Jackson. (*See* Verizon CA Opening Brief, Attachment 1 at Appendix 2 (April 28, 2000).) In this example, if one were to change the value of the highest CLEC result, 5, to 10, the permutation statistic would not change and remains at an alpha of about 0.15 – a "pass" at a critical alpha level of 0.10. In contrast, the Z-statistic would increase considerably, as the CLEC mean would increase from 4.0 to 6.5. The Z-statistic would increase from 1.2 (0.12 alpha) to 3.0 (0.001 alpha), changing this result from a "pass" to a "failure." Generally, non-parametric tests are considered less powerful insofar as they rely on ranked rather than interval data. R. Khazanie, *supra* at 720 (1997).

discrimination. For example, an unusually long delay in obtaining a needed phone service can be especially troubling. Other issues regarding the selection of the Z-test or the permutation test are more fundamental. If it is more appropriate to view the ILEC and CLEC performance results as samples of a theoretically larger process, then the Z-test may be the more appropriate test. If it were more appropriate to view the ILEC and CLEC performance results as the whole population of production output, then the permutation test would be more appropriate. This underlying issue was raised in the ACR, but has not been resolved by the parties or the record in this proceeding. Until we can determine which test is the more appropriate treatment of the data, including underlying issues such as “production output” versus “larger process population sampling” and more specific issues regarding outlier treatment, we are not in a position to either order or approve use of the permutation test. The most important question of decision accuracy is not resolved. Additionally, we need to better understand what the appropriate sample sizes are for using the permutation test versus the Modified Z-test.

Consequently, we will order the Z-test used during the trial period for all average-based performance results. Most importantly, we will not order Pacific to implement a permutation test data analysis system since even the new lower cost estimates warrant a greater confidence than we currently have in the test’s benefits relative to its costs. However, we recognize the permutation test’s potential for being the more accurate test, especially if it is appropriate to view a CLEC result as a sample of a fixed production output result. As we believe it would be a mistake to leave unresolved the questions surrounding this test’s potential, we direct the parties to conduct or fund a research inquiry to answer these questions. We prefer a collaborative research approach where all interested parties would collectively influence the research proposal, and thus would be more inclined to accept the results. But in the interim, the Z-test is the most developed and accepted alternative to permutation testing. We shall order that the Modified Z-test be used for average-based parity performance measures. We discuss further the problem of small samples in a following section.

Percentage-based measures

Modified Z-tests

While the parties have proposed Modified Z-test variants for percentage-based measures, and those variants are being used in New York and Texas, these measures present new difficulties for Modified Z-test application. For example, the test requires an ILEC variance. When there is perfect ILEC performance, the Modified Z-test statistic is not calculable.[101] Pacific proposed a modification to the Modified Z-test for percentages based on the CLEC variance. The CLECs and Verizon CA proposed use of permutation tests, or more specifically, exact tests, which do not require calculation of ILEC variance.

Exact tests

Exact tests are called "exact" because if used consistent with necessary assumptions they calculate the exact probabilities of frequency (counted, rate, proportion) data.[102] They represent a special case of permutation testing. The advantage for our statistical model is two-fold: (1) calculations are

---

[101] Pacific Opening Brief at 9-10 (April 28, 2000).

[102] *See* CLECs' Reply Brief at 12 (May 5, 2000) and 2000 GTE/CLEC Workpaper No. 7: D. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures at 221-225 (1997) (March 30, 2000).

made directly from the raw data, and (2) exact tests have the potential to produce more accurate results for small samples. In the case of the percentage-based performance results data, the Fisher's Exact test is appropriate.[103]

The Fisher's Exact test calculates the probability of an obtained or worse result when the data conform to a two-row by two-column table. Such is the case in the analysis of percentage-based measures where, for example, the first row represents CLEC percentages with the number of "missed dates" for orders in the first column and the actual number of "met dates" in the second column. The second row similarly represents the ILEC data, creating a two-row by two-column data table, or a "2 x 2" table. Given such a table, there is a limited number of possible unique combinations, or permutations, of entries in each of the table's four "cells." The Fisher Exact test determines the probability of each individual combination that is as extreme or worse than the obtained combination being tested. The sum of these probabilities is the probability that the obtained result could occur if the results are only due to random variation.

This probability is "alpha," the probability of a Type I error. Unlike for average-based permutation applications, outliers cannot affect the result, as the data consist only of "cell counts." Additionally, unlike for average-based permutation applications, the results from the percentage-based Modified Z-test and the results from the Fisher's Exact test converge towards equality as theoretically expected.[104] Additionally, the FCC has approved an application that

[103] *Id.* at 221.

[104] We take official notice of sample Fisher's Exact test and Z-test calculations performed collaboratively by staff and Pacific's consultant that show this convergence. The results of these calculations are presented in Appendix D. During the calibration

Footnote continued on next page

uses the Fisher's Exact test for percentage-based measures.[105] We shall order that the Fisher's Exact test be used for all percentage-based parity tests.[106] The evidence before us indicates that it provides accurate decision error probabilities, is consistent with theoretical assumptions, and solves the Z-test application problems.

Rate-based measures

The problem, and our solution, for rate-based performance result analysis is similar to the case of percentage-based performance measures. In this case, a binomial exact test is applied to rate data because the Fisher's Exact test's assumptions are not met. Specifically, the Fisher's Exact test is not appropriate where the row totals are not fixed, or where an entity being observed can contribute more than one cell entry. In the case of percentage-based measures, the Fisher's Exact test is warranted because the row totals are always 100 percent, equal to the total number of CLEC or ILEC orders, and every order only creates one cell entry. In contrast, row totals for rates vary directly with the performance result. For example, the most common rate measure is service "troubles." The rate is typically taken as the rate of troubles per number of lines. This figure can theoretically vary from zero to a number greater than the number of lines

---

phase, parties will be able to confirm these results for the data that is available to them by their own agreements.

[105] *Bell Atlantic New York Order,* 15 FCC Rcd at 4188-4189, App. B., ¶ 13 and n. 39.

[106] Since larger samples cause computer resource problem, an upper sample size limit will be applied. Since Z-test and Fisher's Exact Test have the same results for large samples, and since calculations over approximately 1000 for CLEC "hits" and "misses" can generate computationally difficult numbers, the Z-test will be used for those samples. (See Appendix C.)

because it is possible to have more than one trouble per line. Consequently the row totals are not fixed. However, in this case, assuming the parameters for a Poisson distribution, a binomial exact test can be applied to calculate the probabilities of rate performance results.[107]

Additionally, like the percentage-based Fisher's Exact test applications, and unlike for average-based permutation applications, the results from the rate-based Modified Z-test and the results from the binomial exact test converge towards equality as theoretically expected.[108] Verizon CA, the CLECs, and ORA agree to the appropriateness of the binomial test[109] and Pacific does not object. We shall order that the binomial exact test be used for all rate-based tests as the evidence before us indicates that it provides accurate decision error probabilities, is consistent with theoretical exceptions, solves the Z-test application problems, and is preferred by most parties.

Confidence levels

Alpha levels

The specific fixed alpha levels that have been recommended in this proceeding are 0.15, 0.10, and 0.05 alphas, which correspond to the 85%, 90%,

---

107 CLECs' Reply Brief at 11(May 5, 2000).

108 We take official notice of sample binomial and Z-test calculations performed collaboratively by staff and Pacific's consultant that show this convergence. The results of these calculations are presented in Appendix E. During the calibration phase, parties will be able to confirm these results to the extent that their own agreements allow access to the necessary data.

109 Verizon CA Opening Brief at 24 (April 28, 2000); CLECs' Reply Brief at 11 (May 5, 2000).

and 95% confidence levels, respectively. The 90% confidence level suggested in the ACR is no party's favored level. The ILECs, Pacific and Verizon CA, prefer a 95% level to minimize the possibility of payments made due to sampling error when there are no real differences. The CLECs and ORA prefer an 85% confidence level to minimize the possibility that the ILECs escape payments when there are real differences, but those differences are masked by sampling error.[110] Each side wishes to protect against the negative effect of random variation. But since there are two possible effects of random variation, and as one is minimized the other is maximized, the two sides differ in the preferred confidence level.

Pacific and Verizon CA assert that the 95% level should be used since it is an accepted convention. We disagree. While we understand that it is a convention is some contexts, it is important to understand those contexts to see if they generalize to the present case. They do not. Academic texts that address the use of the 95% level, and that go beyond simply noting its common use as a convention, are clear in pointing out its arbitrariness in applied decision settings:

> *The widespread convention of choosing levels of 0.05 or 0.01 irrespective of the context of the analysis has neither a scientific nor a logical basis.* The choice of level is a question of personal judgment in the Fisherian approach and one of considering type I and II errors in the Neyman-Pearson approach. Since for a given sample size decreasing one error probability

[110] Our conclusion regarding ORA's position here is based on its preference for one standard deviation being the cut-off for a discrimination finding, and its statement describing its position as similar to the CLEC position. ORA Opening Comments on the ACR at 6.) One standard deviation is approximately equivalent to a 15% alpha, or an 85% confidence level. However, as discussed *supra,* ORA's definition of "standard deviation" is unclear.

> increases the other…, it is possible to argue for a relative balance. In particular, if at $\alpha = 0.05$ the power is very low, one might seriously consider increasing α and so increasing the power.[111]

> In our opinion, there is no "right" or "wrong" level here – the decision must be made in full consideration of parameters inherent in the problem itself. It is doubtful that setting *a priori* levels of .05, .01, or what have you settles the matter.[112]

> No absolute standard can be set up for determining the appropriate level of significance and power that a test should have. The level of significance used in making statistical tests should be gauged in part by the power of practically important alternative hypotheses at varying levels of significance. If experiments were conducted in the best of all possible worlds, the design of the experiment would provide adequate power for any predetermined level of significance that the experimenter were to set. However, experiments are conducted under the conditions that exist within the world in which one lives. What is needed to attain the demands of the well-designed experiment may not be realized. The experimenter must be satisfied with the best design feasible within the restrictions imposed by the working conditions. *The frequent use of the .05 and .01 levels of significance is a matter of a convention having little scientific or logical basis. When the power of tests is likely to be low under these levels of significance, and when type 1 and type 2 errors are of approximately equal importance, the .30 and .20 levels of significance may be more appropriate than the .05 and .01 levels.* (p. 14, emphasis added.)[113]

---

111 A.H Welsh, Aspects of statistical inference at 128, (emphasis added) (1996).

112 J. Skipper, A. Guenther & G. Nass, *The sacredness of .05: A note concerning the uses of statistical levels of significance in social science*, 2 The American Sociologist at 17 (1970).

113 B.J. Winer, *supra* (1971).

> In principle, if it is very costly to make an error of Type II by overlooking a true departure from [the null hypothesis] but not very costly to make a Type I error by rejecting [the null hypothesis] falsely, one could (and perhaps should) make the test more powerful by setting the value of [alpha] at .10, .20, or more. This ordinarily is not done in social or behavioral science research, however. There are at least two reasons why [alpha] seldom is taken to be greater than .05: In the first place. . . in such research the problem of relative losses incurred by making the two kinds of errors is seldom addressed; hence conventions about the size of [alpha] are adopted and [beta] usually is ignored. The other important reason is that given some fixed [alpha], the power of the test can be increased either by increasing sample size or by reducing the standard error of the test statistic in some other way, such as reducing variability through experimental controls. (P. 290.)[114]

These four quotes point out the dilemma in our applied problem. Unlike in scientific applications where the parameters of an experiment are easily manipulated, we have neither the luxury nor the discretion to change the sample size, the effect size, or the sampling error. Consequently, the Commission must chose an alpha level without regard for conventions developed in qualitatively different contexts.[115]

Additionally, while the authors of the last two quotes appear to differ in their recommendations regarding the relative consequences of Type I versus Type II error, these differences should be viewed in terms of different

---

[114] W. Hays, *supra* (1994).

[115] Faced with a similar problem in D.97-09-045, we based our decision on the actual probabilities, and not on an arbitrary pre-selected significance level. (D.97-09-045, *mimeo.* at 31-32 (September 3, 1997).)

assumptions regarding the freedom to change sample sizes, error terms, and the strength of experimental treatments, among other parameters. Academic treatises directly addressing these relative consequences have developed formulas that balance the net consequences of any resultant error by establishing loss functions. [116]

For example, while different alpha, and thus beta, levels are appropriate depending on the ratio of the costs of the consequences of both types of errors, when the error consequences are deemed to be equal, losses are minimized when alpha and beta are set to be equal.[117] We have not determined a specific ratio for the relative consequences of failing to identify competition barriers when they exist versus monetary payments made when they should not be made. However, at this point we can only assume from the purpose of the Act and the regulatory policy mandating competition,[118] that the consequences of not identifying barriers is at least equal to any misappropriated payments.[119] As

---

[116] C. Das, *Decision making by classical test procedures using an optimal level of significance*, 73 European Journal of Operational Research at 76-84 (1994).

[117] *Id.* at 78.

[118] For example, the FCC has stated that it based its public interest evaluation and approval of BANY's 271 application on the fact that a primary element of the New York remedies plan was its design to "detect" discrimination. FCC BANY Order at ¶ 429. Test power is the closest index of this fundamental purpose.

[119] In comments on the draft decision, the ILECs dispute that failure to detect discrimination has consequences as harmful as mistakenly detecting discrimination. Pacific Comments at 8-9 (December 18, 2000); Verizon CA Comments at 9 (December 18, 2000. We agree that this issue deserves further discussion, but we are also comfortable moving forward with an interim decision based on our assumption. There will be ample opportunity to further consider these issues before any element of the

Footnote continued on next page

a consequence, our goal will be to choose an alpha level that serves to balance with a beta level.[120] In doing so we are not addressing risk. The question of relative risk is more appropriately addressed in the proceeding's next phase, which will establish the "consequences" for the performance decisions made in the present phase. Balancing alpha and beta to be equal only ensures that the most accurate decision is made, not what the consequences of those decisions will be.

We note that the FCC encourages such a balance.[121] We also note that the NYPSC has adopted as low as an 80% confidence level in certain circumstances, possibly to achieve a better balance. While we have discussed a 90% confidence level as a compromise to facilitate negotiation progress, we are unwilling to permanently select such a fixed level based solely on the midpoint between two negotiating positions.

Pacific argues against the 90% confidence level stating, "There is no forum of which we are aware that supports the use of a 10% error rate." However, we find it notable that the BANY remedies plan uses a 21% error rate (79% confidence level) for conditional failure identifications and what in essence

model that depends on this assumption is implemented and before a final remedies plan is implemented.

[120] The parties have argued over balancing for "equal risk" versus "equal error." (*E.g.*, Verizon CA Reply Brief at 9 (May 5, 2000) We note that when the ratio of error consequences is set to "1," the Das (1994) "equal risk" formula simplifies to what essentially is an "equal error" formula.

[121] *Bell Atlantic New York Order*, 15 FCC Rcd at 4190-4194, App. B., n. 50.

is a 10% error rate for final disposition of those identifications.[122] We also note that one of the statistical texts frequently cited in the FCC's BANY 271 approval states, "The value of alpha chosen is usually between 0.01 and 0.1, the most common value being 0.05."[123]

Although Verizon CA presents an academic cite as justification for its preference for a 95% level (.05 alpha), we find that that cite refers only to less formal "rough conventions" and does not refer to any context or consequences of the two different types of error.[124] Additionally, Verizon CA quotes an affidavit in a FCC proceeding citing an AT&T statistician's support for the 95% level. We also do not find that quote necessarily applicable to the problem of balancing the two errors. In that quote, Dr. Mallows states that a 95% level would control Type I error "while making the probability of Type II errors small for violations that are of substantial size."

The Commission cannot base its decision on such a statement when the statement context is not clear. At the time Dr. Mallows made the statement, over two years ago, it may not have been apparent how small the sample sizes

---

[122] *Id*, at 4189, App. B, n. 41.

[123] Khazanie, *supra*, at 506 (1997).

[124] "The hypothesis test of H_0 consists of computing [the achieved significance level], and seeing if it is too small according to conventional thresholds. Formally, we choose a small probability α, like .05 or .01, and *reject* H_0 if [the achieved significance level] is less than α.... Less formally, we observe [the achieved significance level] and rate the evidence against H_0 according to the following rough conventions: [achieved significance level < .10 [is] borderline evidence against H_0." B. Efron & R. Tibshirani, An introduction to the bootstrap at 203-204 (1993) (emphasis added).

were going to be, and thus he may have been referring only to results obtained from fairly large samples. We are concerned that even substantial Type II errors may not be identified with a 0.05 alpha level for small-to-moderate samples. Additionally, Dr. Mallow's statement implied that the statistical test, through its significance level, was used to determine magnitude as well as statistical significance. We cannot know how Dr. Mallows' statement applies to our context without knowing what he meant by the term "substantial." Dr. Mallows more recently has stated that he believes 0.15 is the appropriate level and that the 0.05 level seems too small since it "gives more of a chance of failing to detect a violation than of performing a Type 1 error. . ." (RT at 919, lines 14 to 24) But more importantly, our approach is different. We will address the magnitude issue separately below after the error problem has been addressed.

A deciding factor for us is the potential consequences of the two types of error to our overall performance remedies plan. Given the potential for us to err on one side where we might favor either alpha levels or beta levels to the detriment of the other, the correctability of any such imbalance that might result is an important consideration. On one hand, if we set alpha too large and as a result make Type I errors, we can make up for these errors in the incentive-amount methodology phase of this proceeding. For example, we could adjust the incentive amount to the actual Type I error calculated for each performance result. Specifically, presented for illustration purposes only, we could levy an incentive payment for a result with a Type I error probability of 0.05 at 95% of a pre-determined amount, but levy a payment with a Type I error probability of

0.15 at 85% of the same amount.[125] In contrast, once we have made a Type II error, no correction is possible since parity would have been concluded. In this case the measurement would not make it to the incentive payment phase, and thus would not be correctable.[126]

We note that the NYPSC addressed this issue by selecting three alpha levels: a 0.05 alpha level for immediate non-parity identification, approximately a 0.20 alpha level for conditional parity identifications depending on subsequent months' results, and what in essence is a 0.10 alpha level for final disposition of conditional identifications.[127] The parties have variously proposed

[125] The actual alpha probability for each result would be used, not any pre-selected alpha level. For example, if the probability of an obtained result being a Type I error was .03, then 97% of the payment would be assessed, if the error was .12, then 88% of the payment would be assessed, and so forth. Across time, this method may mitigate the problem of Type I error payments. For example, in the long run, there may be no difference between "forgiving" 15% of the incentive payments versus charging only 85% of the levied payments. A probability-adjusted scheme would be even more accurate in the long run. *See* H. Raiffa, Decision analysis (1970). We provide this example for illustrative purposes only and do not suggest that these values would be the specific appropriate ones. Our point is that payments can be scaled to error probability estimates similar to that suggested in the ACR. ACR at 26 (November 22, 1999).

[126] In comments on the draft decision, Pacific disputes our conclusion. Pacific Comments at 7-8 (December 18, 2000). We appreciate its comments and welcome its interest in understanding Type II errors. However, we are not persuaded by its argument. Pacific appears to be discussing a different topic – the likelihood of future discrimination being detected. Our point is that nothing can be done about an erroneous decision to conclude parity, because assumed parity causes no action or adjustments regardless of the degree of the error. On the other hand, when it is concluded that discrimination exists, the degree of Type I error probability is apparent, action is taken, and that action can be "calibrated" to the degree of the error.

[127] *Bell Atlantic New York Order*, 15 FCC Rcd at 4189, n. 41.

the 0.05 or the 0.15 alpha levels, and the ACR recommended a 0.10 level for the purposes of development, inquiry and compromise. However, we are not comfortable selecting alpha levels without discussing and assessing beta and its converse, test power.

Test power

Unfortunately, the record is relatively silent on the actual beta values that various critical alpha levels might produce. The only estimates in the record are that in early tests, AT&T estimated betas to range as high as 0.21 when critical alpha levels were set to 0.05.[128] A beta value of 0.21 corresponds to a test power of 0.79, or 79%. AT&T also estimated that if alpha was set to 0.15, then betas would average a similar level - an average test power of 85% when the average Type I confidence level is 85%. Yet it is unclear if the results from the earlier tests are comparable to the performance results in California. To remedy this lack of critical information, we shall direct the ILECs to calculate both alpha and beta values whenever a statistical test is applied.

Staff has performed some preliminary estimates of beta values using four different alpha levels.[129] The results are discouraging about the ability of our

---

128 Verizon CA Reply Brief at 8, n.2 (May 5, 2000).

129 We take official notice of tables prepared by staff summarizing the beta levels that are obtained with different tests and different alpha levels. These tables are presented in Appendix F. These values are based on May 2000 performance data and are preliminary estimates based on the application of the Modified Z-test to average, percentage, and rate-based measures. The alternative hypotheses posed for all estimates were that the CLEC's results were at least 50 or 100 percent worse than the ILEC's results. The formula used is found in Hays, *supra* at 284-289 (1994). Staff presents these values as approximations, and does not represent that these calculations are necessarily the best estimate of beta. We present them here to begin a discussion of

Footnote continued on next page

model to perform its most fundamental task, to detect competition barriers. For example, with a 0.10 critical alpha level, and selecting a 50 percent difference to establish alternate hypotheses, beta values average 0.63 with a median of 0.79.[130] While the selection of a 0.10 critical alpha threshold ensures that 100 percent of the performance results are subject to a 10 percent maximum Type I error, it only provides that 16 percent of the results are subject to a 10 percent maximum Type II error.[131]

Additionally, the parties have not recommended any minimum test power, or its respective error, beta. Since beta is determined by the other elements, the degree of test power ends up being that which results from the other elements. The record is relatively silent on the appropriate test power or beta error level. While unfortunate, this state of affairs is understandable since at the outset alpha can be set, but beta can only be determined upon obtaining the measured performance results. Beta will thus vary for every performance result. For every obtained result, however, it is possible to balance alpha and beta if we can safely make assumptions about two components of the analysis: (1) the relative consequences for each type of error, and (2) the specification of the alternative hypothesis.

As a general policy statement, it is reasonable to assume that a Type II error is at least as important as a Type I error, as discussed earlier. Apparent

---

beta estimation, and believe that the values are sufficiently appropriate for us to base the decisions we make regarding the need for further research and development.

130 App. F at 2.

131 *Id.*

discrepancies can be adjusted in the incentive payment phase. However, specification of an alternative hypothesis is more difficult. The alternative hypothesis is the hypothesis that barriers exist - that ILEC service to its own customers is actually worse than to CLEC customers beyond that which could be explained by sampling error. We are aware of three ways to specify the alternative hypotheses. First, the critical value for the alternative hypothesis could be set to equal the critical alpha level value. This would not be much help because the beta error level would always be 50%.

Second, the actual result could be selected as the alternative hypothesis. It would be reasonable to assume that an actual result was the best estimate of the actual underlying process, and as such best represents the alternative hypothesis. A statistical test could then estimate the respective Type I and II errors of this result being a "true" mean, not identified due to sampling error. In this case, the balanced alpha and beta level could easily be determined.[132] It is unclear at this point, though, what the effects of this balancing would be since for very small differences, both beta and alpha might be very large, whereas for big differences, both might be small. If this happens, we would still have to set some alpha / beta thresholds, and / or set some "material" difference thresholds.

Third, the critical alternative hypothesis value could be determined by identifying a performance result or level where ILEC and CLEC service differences become "meaningful." Verizon CA has proposed a few performance

---

[132] C. Das, *supra*, at 78 (1994).

levels, called "deltas," as a solution to a different problem in this proceeding.[133] However, the record contains no information on what most deltas would be, as no party has submitted any proposal containing a comprehensive set of specific deltas.

In comments on the draft decision, the ILECs assert that establishing alternate hypotheses that represent competition barrier thresholds is a significant problem that may make the exercise fail. Pacific Comments at 9-12 (December 18, 2000); Verizon CA Reply at 4 (December 22, 2000). However, we note that the ILECs are willing to establish nearly identical thresholds to use for "materiality" standards to reduce payments. Pacific Comments at 20 (December 18, 2000); Verizon CA Comments at 11- 12 (December 18, 2000). Adapting such standards for alternate hypotheses, if any adaptation is necessary at all, should be relatively easy.

A fixed alpha is not an adequate long-term solution. As the CLECs have asserted and as staff's data analysis has shown, test power is very low for the small samples that represent the majority of the performance measure results. On the other hand, the ILECs have asserted, and staff's data analysis confirms, that fixed alphas that provide better test power for small samples result in unnecessarily high test power for large samples. This unnecessarily high test-power can easily result in meaningless differences being found statistically significant.[134] We believe that the problems of insufficient test power for small

---

[133] Verizon CA Opening Brief at 10-11 (April 28, 2000).

[134] Verizon CA Opening Brief at 10, n. 6, citing P. Bickel & K. Doksum, Mathematical statistics: Basic ideas and selected topics at 175 (1977)(April 28, 2000).

samples (large beta) and "too much" test power for large samples can be better resolved through even approximate alpha / beta balancing techniques. We direct the parties to develop and implement an alpha / beta balancing procedure for our model. However, to give sufficient time for its development without delaying Pacific's 271 application, we shall adopt a fixed alpha solely for the interim, and shall order that the balancing components to the model be added by the end of the trial period unless the parties reach agreement and move to implement the components sooner.

Fixed alpha

We conclude for the reasons cited above that a fixed alpha critical value should only be used as an interim decision-criterion solution. Setting alpha to remedy one problem only makes another. We select a larger alpha level, 0.10, instead of the 0.05 level to enhance decision accuracy and to avoid uncorrectable decision-making errors while still being able to address correctable errors in the next phase of this proceeding. We select a smaller alpha level than 0.15 because we are concerned about the effect on large-sample results. We have selected the 90% confidence level (0.10 alpha, or 10% significance level) to control the Type I error and to reduce the Type II error to more acceptable levels for the preponderance of the performance results. That is, we choose to be at least 90% confident that any barriers we identify represent real differences, not differences due to sampling error (random variation), while increasing the average confidence level (power) for detection of actual differences from 30% for the 0.05 alpha to 37% for the 0.10 alpha.[135]

Additionally, because of the low power of these tests, pending further development and consideration we intend to also adopt the 80% confidence level (0.20 alpha) for conditional failure identifications. This threshold is used in the BANY performance remedies plan for conditional identifications where results at 0.20 alpha or less were deemed failures if they occurred in two months of a three-month period.[136] We will not dictate the additional specifications for such conditional identifications, but instead direct parties to set forth those specifications in the next phase. Among other possibilities, our plan could have additional criteria such as (1) successive failures such as in the BANY plan, (2) alpha and beta balance at values less than 0.20, or (3) for CLEC-specific performance assessment, industry aggregate performance out of parity. Noting that if a 80% confidence level (0.20 alpha) was used as the overall fixed threshold instead of the 90% level (0.10 alpha), average

---

[135] These figures are based on an alternate hypothesis of 50% worse performance for the CLEC and on CLEC samples of only 5 or more. Average power increases from 37% to 49% assuming a 100% worse-performance alternate hypothesis. These estimates were made from existing data and could easily change in the future without any changes in the plan. For example, if the CLECs gain larger shares of the local phone markets and CLEC companies individually place more orders, sample sizes will increase, with a resulting increase in test power, with all other elements held constant.

[136] *Bell Atlantic New York Order*, 15 FCC Rcd at 4189, App. B, n. 41

power would increase from 37% to 48%,[137] we wish to take advantage of this increased power at least on a conditional basis.[138]

Material differences

None of the parties have specified a comprehensive set of minimum differences (effect size) between the ILEC and CLEC performance results that would identify a competition barrier. Two parties have raised the issue. AT&T has somewhat tangentially raised the issue in its discussion of test power[139]. To calculate test power, an alternative hypothesis must be specified as discussed *supra*. AT&T estimated test power across an array of different performance results after subject matter experts made judgments creating competition-affecting performance thresholds.[140] Verizon CA currently proposes utilizing "deltas" which embody virtually the same concept, albeit for different

---

[137] These figures also are based on an alternate hypothesis of 50% worse performance for the CLEC and on CLEC samples of only 5 or more. Average power increases from 48% to 60% assuming a 100% worse performance alternate hypothesis. See Appendix F.

[138] Some commenters raise concerns that the 0.20 alpha level was not addressed in the record. Verizon CA Comments at 12-13 (December 18, 2000). Pacific Comments at 13 (December 18, 2000). However, we note that there was considerable discussion of what the appropriate alpha level would be and at least one party speculated without the benefit of current data that the level would be 0.15. The record has sufficient discussion of appropriate alpha levels for us to order further development on optimal levels, such as the 0.20 level. We advise parties that we cannot guarantee any conditional 0.20 level will be adopted in the final model, especially if we find that no party has specified reasonable conditions for implementing this alpha level. Nevertheless, we note that a closer read of the New York Public Service Commission's use of this alpha level is likely to be informative.

[139] Cited in Verizon CA Reply Brief at 8, n. 2 (May 5, 2000).

[140] Id.

purposes.[141] Whereas AT&T created thresholds to investigate insufficient test power, Verizon CA proposes to create these conceptually identical thresholds to investigate "too much" test power.[142] We find that both efforts to establish "material" thresholds have merit. First, as we have described above, test power is a primary decision-accuracy concern for this remedies plan. The best way to calculate test power is to specify a meaningful alternative hypothesis, and the most meaningful alternative hypothesis is one that embodies the core performance remedies plan goal, barriers to competition. Second, it would be contrary to the same decision accuracy policy goals to impose incentive payments when an ILEC is providing virtually the same service to a CLEC that it is providing to itself with no negative impact on competition. Recent academic discussions have pointed out that in the case of large samples, statistical results right at an alpha level of 0.05, for example, can provide evidence *for* the null hypothesis, rather than against it as designed:

> Results indicate that for point null hypotheses, a statement of [statistical significance at alpha] does not have a straightforward, evidential interpretation. It is demonstrated, that for larger samples particularly, that a report merely that data are [statistically significant at alpha] has no objective

---

141 Verizon has proposed specific "deltas" for nine of the approximately thirty performance measures that will be included in the performance incentives plan, although Verizon calls these proposals "preliminary values" that "can and should be adjusted as more data is gathered." Verizon CA Comments on Workshop at 9-11 and Attachments 2 and 3 (April 28, 2000). We appreciate this explicit proposal and encourage further development. Notably, no party has proposed material differences for a majority of performance measures or any average-based parity measure.

142 Verizon CA Opening Comments on ACR at 11 and App. B. at B2 - B3 (January 7, 2000); Verizon CA Reply Brief at 8, n. 2 (May 5, 2000).

> meaning, and under some conditions should be interpreted not as evidence against the null hypothesis, as is usually supposed, but as strong evidence in its favor.[143]

For very large samples, significant differences at or close to the .05 threshold might be so negligible as to be perceptually the same to a CLEC customer as would be the "statistically significantly different" ILEC service, and as a consequence actually be evidence of parity, not discrimination. Statisticians seem to agree that statistical significance is different from substantial significance.[144]

We find that the "material difference" standard has merit and the potential to improve the decision model we specify. However, we are concerned that the task to construct a set of difference thresholds is difficult, and yet to be accomplished in any collaborative forum. We encourage the parties to complete this task as part of the alpha/beta balancing task we order today. However, since other ways to specify an alternative hypothesis may be easier to accomplish, yet sufficient to enhance decision accuracy, we will not order the material differences be defined for every measure. Other methods for balancing alpha and beta errors may resolve the material difference versus statistical difference problem and we choose to allow the parties the discretion to collaboratively determine the best solution before we order our own solution.

---

[143] D. Johnstone & D. Lindley, *Bayesian inference given data "significant at α": Tests of point hypothesis,* 38 Theory & Decision at 51 (1995).

[144] For example, *see* D. Gold, *Statistical tests and substantive significance,* 4 The American Sociologist at 42 – 46 (1969).

Optimal alpha and beta levels

The parties have variously discussed “equal risk,” “equal error,” and “balancing alpha and beta.” “Equal risk” refers to a situation where the expected consequences of the performance remedies plan are the same for an ILEC as for the CLECs. The concept of equal risk is beyond the scope of our decision model as it necessarily requires incentive payment specification which we will not consider until the next phase of this proceeding. “Equal error” and “balancing alpha and beta” refer to a situation where the two possible decision-making error probabilities are the same. We endorse the concept not only because it meets our fairness principle, but also because it maximizes decision accuracy.

Overall decision error is minimized when alpha and beta are balanced.[145] But most importantly, if we are to create a “level playing field,” we must be fair in our acceptance of decision error. The data shows that a fixed alpha level of 0.10 can only be suitable for an interim implementation because it favors reducing the error that only the ILECs wish to reduce. There would be no level playing field if we tolerated no more than 10 percent error harmful to the ILECs, yet tolerated 40 to 60 percent error harmful to the CLECs. We only take the 10 percent alpha level as an interim compromise necessary for progress. Additionally, maximizing decision accuracy by equating possible errors is an appropriate first step to optimizing equal risk, and does not necessarily interfere with the consequence-setting function of the next phase of this proceeding.

---

[145] C. Das, *supra* (1994).

In comments to the draft decision the ILECs raise several arguments against attempting to balance alpha and beta error. Pacific Comments at 9-12 (December 18, 2000); Verizon Comments at 9-11 (December 18, 2000). We take these comments seriously and note that there will be time for further discussion and consideration of these issues before, and if, we decide to implement a executable balancing feature. However, we note that neither ILEC acknowledges the balancing plan’s potential benefit of lowering the average Type I error and of reducing small magnitude failure identifications.[146] We are optimistic that parties will find the net result of an error balancing plan mutually beneficial once the details are resolved.

We direct the parties to work collaboratively to develop and implement an alpha / beta balancing decision component for our decision model by the end of the trial period. If the parties are unable to agree on such a model component at that time, we shall direct parties to submit their individual models for our review and decision.

---

[146] We note that balancing alpha and beta levels can be a “win-win” situation for the parties when compared to a fixed alpha level. Examining CLEC samples, staff has noted that whereas a fixed alpha of 0.10 results in a maximum error rate of 0.10 for all analyses, if alpha and beta are balanced and the maximum error rate is raised to 0.25 for all analyses, the resulting *average* maximum error rate is 0.072, well below the 0.10 fixed alpha level. A summary of staff’s analysis is provided in Appendix G, which also shows the error balancing effect to reduce detection of small differences and increase the detection of large differences.

Minimum sample size

Minimum sample size requirements vary depending upon the type of statistical test used. For example, as discussed above, exact tests are not dependent on inferences about the underlying distribution, therefore the accuracy of calculated alphas is relatively unaffected by sample size. Therefore we find it necessary to discuss sample size issues individually for each type of measure.

Average-based measures

Sample size requirements for average-based measures are the most difficult to resolve. On one hand, the CLECs have pointed out the importance of separately assessing performance for even the smallest CLEC with the least activity since these CLECs depend more on each order or service than do the larger CLECs. Harmful ILEC performance in small new or innovative market niches, or harmful ILEC performance to smaller CLECs could be masked by larger market samples or larger CLEC samples when the results for CLECs are combined ("aggregated"). If so, then the smaller markets and the smaller CLECs would not be provided the protection that this performance remedies plan is supposed to provide. Such small CLECs and markets effectively would be unprotected by competitive market reforms, and thus might fail.

Consequently, the CLECs have urged sample sizes small enough to protect these markets. We agree with this principle, and thus, one goal of our plan is to assess each CLEC's performance results for each submeasure. On the other hand, as sample sizes become small, Central Limit Theorem states that the normality desired for Z-tests can no longer be assumed. The accuracy of the error estimates, alpha and beta, becomes suspect with the smaller samples. So we are faced with the potential dilemma of having to choose between achieving greater decision accuracy or protecting an important sector of the market. The parties predictably were not able to agree on a solution to this dilemma. Proposals ranged from a sample size minimum of 1 to a minimum of 50 or more.

The issue is relatively simple for the ILECs. They are concerned that small samples could produce inaccurate error estimation, which could inappropriately subject them to payments even when their processes are non-discriminatory. However, since the ILECs are more concerned with alpha levels, unlike beta levels, alpha levels can be held constant regardless of the size of the sample. So even though there may be an issue of accurate alpha estimation, there is still some adjustment as sample sizes decrease – alpha error is held constant. Additionally, with alpha error held constant and as sample size decreases, test power decreases, thus reducing the ILEC's potential liability under any performance remedy payment plan. On the other hand, the ILECs may be concerned that smaller samples generate greater incentive payment exposure by the consequent that there are more performance tests. However, this concern is best addressed in the incentive payment phase where it can be accommodated if warranted. The ILECs also prefer aggregation of all results, since in their view, the total result is the best indicator of the parity of the process.[147] As a compromise, the ILECs offered to use sample sizes from 5 to 20, and they have offered to aggregate results in order to achieve these minimum numbers. With a few exceptions, the ILECs wish to exclude, from the performance remedies plan, data that does not meet these sample minimums.[148]

---

147 Pacific Reply Brief at 11 (May 5, 2000).

148 The ILECs and CLECs have agreed to have no minimum sample size requirements for "rare submeasures," which are submeasures that rarely see activity, yet are so

Footnote continued on next page

For example, samples that contain four or less observations after aggregation rules have been applied would be discarded unless they are a designated "rare submeasure" that should be analyzed regardless of sample size.

The issues for the CLECs are more complicated. On one hand, since increasing the sample size increases test power as the significance level is held constant, the CLECs would seem to prefer larger samples. Smaller samples often have negligible test power. However, on the other hand, the CLECs prefer no aggregation of results since the actual service each company receives is critical to them. Each company is directly affected by the service it receives from the ILEC independently of the service that other CLECs receive. Consequently, the CLECs have urged inclusion of sample sizes small enough to protect these markets. Second, the CLECs urge that all data be analyzed regardless of sample size. They do not want any data discarded from the performance remedies plan. It is unacceptable to the CLECs to ignore poor performance to a small emerging CLEC, simply because of a minimum sample size rule. However, like the ILECs, the CLECs agreed to a compromise position, accepting some aggregation rules, but firmly rejecting exclusion of any performance results because of insufficient sample size.[149]

Assisted by Pacific's technical expert, staff examined how one possible compromise set of aggregation rules would function.[150] In summary,

important as to need close monitoring when any activity occurs. These submeasures are listed in Appendix H, Attachment 1.

149 CLEC's Reply Brief at 8-9 (May 5, 2000).

150 We take official notice of staff's summary of this analysis, included here as Appendix H.

the rules were as follows: (1) Samples of 10 or more would be separately analyzed; (2) All samples of less than 10 would be aggregated for a collective analysis if they achieved at least a sample size of 5; (3) Where a minimum of 5 was not achieved, the remaining samples would be aggregated for analysis with all other CLECs for the submeasure; and (4) Where the industry aggregate did not achieve a minimum of 5 the data would be discarded.[151] Using these rules, for the most recent month presented, March 2000, 57 percent of the performance results could be analyzed without aggregation, 39 percent could be aggregated with other small sample results, 1.3 percent had to be aggregated with the rest of the industry, and 2.4 percent of the results had to be discarded.[152] While not having an opportunity to comment on this, the CLECs can be anticipated to object to these rules insofar as they require that 43 percent of the results be aggregated or discarded and that 3.7 percent (127) be either aggregated with the whole industry, possibly rendering their results masked by a much larger sample, or be discarded.[153]

Staff found several unresolved problems with the proposed compromise aggregation rules. First, in some cases, even with very low test power for a reasonable alternative hypothesis,[154] the performance results to a small CLEC were highly statistically significant with an extremely low Type I

---

[151] Pacific's Reply Brief at 11(May 5, 2000).

[152] *Id.* at 12.

[153] CLEC Reply Brief at 2 (May 5, 2000).

[154] The alternative hypothesis was that performance for CLEC customers is at least 50% worse than for ILEC customers.

error, or alpha. However, the aggregation rules caused this result to be combined with and masked by results for large CLECs. Second, in other cases, where several small CLECs experienced better or nearly equal ILEC performance, exceptionally poor performance to one CLEC caused the aggregate performance to be identified as a failure. Such an outcome could trigger payments to each of the CLECs, thus spuriously expanding the ILEC's liability.

Third, the aggregation rules caused some unnecessary aggregation. For some submeasures where only one CLEC did not have the minimum of five or ten results, its results were aggregated across the entire CLEC industry, which often had more than a thousand individual performance results. This would occur even though aggregating with only the smallest CLEC result over five or ten would have provided a sufficient sample size. With the proposed rules the small CLEC result was unnecessarily completely masked by the very large CLEC samples.

Fourth, in cases where there are multiple results for the same CLECs it is not clear which result would be used. For example, when a small CLEC's results are aggregated with larger CLECs' sample sizes that are small, but which are big enough to be analyzed on their own, two different conclusions could be reached. When the larger individual sample results all pass and when the combination of these results do not pass, the individual larger samples will be deemed to have passed individually but not in the aggregate. This result poses a dilemma in that on one hand the aggregate may be the better indicator of the larger process if one assumes a "process model," but on the other hand, assuming a "service model," only the smallest CLEC suffered harm. Each assumption suggests a different remedy.

We believe that it is important to examine performance at the smaller market and smaller CLEC levels. This market arena may be critical for entry and innovation, which in turn are critical to a healthy competitive telecommunications infrastructure. However, given the unresolved issues for sample size and aggregation rules, and the fact that the rules for incentive payments are integrated with the aggregation rules, we are reluctant to permanently order any minimum sample sizes because any such minimums would require some data be discarded. Before finishing this discussion, we examine proposals that might not require sample size minimums.

Permutation testing has been proposed as a solution to the Z-test's small sample normality assumption violations. We prefer use of the permutation test rather than the complicated, and somewhat confusing, data elimination and aggregation rules. However, as we discussed earlier, the record is not sufficiently complete for us to be confident that permutation testing is free of other problems. In New York, while permutation testing is being developed, the New York Public Service Commission has ordered *t*-tests used for small samples as an interim solution for the Z-test small sample problem.[155]

Statistical texts indicate that the *t*-distribution is more appropriate for tests between two sample means, especially for small samples.[156] Use of a *t*-

---

[155] *Bell Atlantic New York Order,* 15 FCC Rcd at 4187, App. B., ¶ 11.

[156] For small samples the distribution of the means of samples is different from the distribution of the raw scores themselves as expressed in Z-tables. Roughly speaking, the mean sample distribution is narrower and taller in these circumstances than the raw score distribution. Consequently, a *t*-distribution should be used for statistical comparisons of means from smaller samples.

distribution "look-up" table could alleviate some ILEC concerns regarding possible alpha estimation inaccuracy for small samples. For example, with the current fixed critical-Z decision rules, a Modified Z-test statistic of 1.8 would identify a failure at all parties' favored alpha levels since it exceeds the most conservative proposed critical value of 1.645. This result would be the same for all sample sizes including a sample size of one. However, the ILEC's concerns regarding alpha accuracy increase as the sample size decreases. Using the *t*-distribution table would adjust for decreasing sample size. For example, for an ILEC sample size of two (*df* = 1), a critical value of 3.078 must be exceeded for the 0.10 alpha level.

Our example of a Z-statistic of 1.8 would not be significant unless the result sample size was at least four, since the critical *t* for a sample of 3 (*df* = 2) is 1.886 and the critical *t* for a sample of 4 (*df* = 3) is 1.638.[157] Consistent with the academic justification of the Modified Z-test, we shall order the test statistic compared to the *t*-distribution. In this regard, we will refer to the Modified Z-test hereinafter as the Modified *t*-test, also consistent with its academic reference.[158]

Unfortunately however, this adjustment affects only the relatively infrequent small ILEC samples and not the preponderance of small CLEC samples.[159] Additionally, other questions still remain regarding the accuracy of

---

157 This illustration uses the ILEC sample size for "looking up" the critical *t*-statistic distribution value. The Brownie, et al., *supra*, research indicates the ILEC sample size should be used for the "lookup" step.

158 Brownie, et al., *supra* (1990).

159 *Id.*

alpha estimation even with more conservative t-distribution tables. Even though the *t*-distribution is a remedy for small samples, its appropriate use still assumes the population is normally distributed, especially for one-tailed tests.[160]

We find that the controversies over the appropriate minimum sample size involve several unresolved elements of our decision model: alpha estimation accuracy, permutation or Modified Z-test use, aggregation rules, data exclusion rules, and incentive payment rules. For the reasons that there are several possible solutions to the minimum sample size problem, the resolution of any one of these problems may resolve the others, and the ultimate solution may necessarily involve decisions about incentive payment rules, we are reluctant to order a permanent minimum sample size. We are concerned that without further information, research, and calibration information, we would be essentially deciding "in the dark." While we prefer not to delay specifying final model components, in this case the complexity of the problem and the potential for a better solution warrants the delay. A better solution may be achieved during the calibration phase when parties can see how various rules, tests, and distributions work.[161]

However, we also are concerned that the parties may not either create or agree on a better solution to the small sample size problem. If this turns

---

[160] Hays, *supra* (1994) at 327-328.

[161] Even in the unlikely event that parties are unable to resolve the small sample problem in the incentive phase, Pacific will still be able to present a completed performance remedies plan to the FCC, either as the "no minimum" default we order today, or a different sample size plan that we may subsequently order for a completed remedies plan.

out to be the case, then we would in effect be ordering many applications of statistical analyses and decision rules for samples as small as one or two individual performance results. We find that we need to set some minimal rules that, in the case that parties are unable to agree on better solutions, will reduce dependence on such very small samples. We shall order the following rules as an interim solution as a "floor" for sample sizes. These rules are designed to avoid discarding any data, and to increase sample sizes for the very smallest samples with minimal impact on the actual results. These rules are also designed to be easily understood with the results easily reproduced. We find that the previously proposed rules are complicated and fall short of our goal of simplicity.

The following rules shall be used for average-based parity performance measures:[162]

(1) For each submeasure, all samples with one to four cases will be aggregated with each other.

(2) Statistical analyses and decision rules will be applied to determine performance subject to the performance remedies plan for all samples after the aggregation in step (1), regardless of sample size. For example, if samples with as few as one case remain after the aggregation, statistical analysis and decision rules will be applied to determine performance subject to the performance remedies plan to these samples, just as they are for larger samples.

These small sample aggregation rules minimize most of the problems described above for Pacific's proposed plan. (See Appendix I.) We do not presuppose how payments will be triggered or allocated under these

[162] The results of these aggregation rules are illustrated in Appendix I.

aggregation rules. The issues will be addressed in the upcoming incentives phase. For example, the parties can decide whether any CLEC whose results are aggregated into a failing aggregate, yet whose individual results are better than the ILEC parity standard, should receive incentive payments. In this case, an "underlying process" model might suggest that this CLEC receive payment because the process was flawed and the incentive was necessary to motivate process improvement. On the other hand, a "service" model might suggest that this CLEC not receive payment since it suffered no competitive harm. In comments to the draft decision, the ILECs seem particularly concerned about assessing small samples for potential remedy payments. Pacific Comments at 16 (December 18, 2000); Verizon CA Comments at 14 (December 18, 2000). We remain receptive to the proposal that any penalty amounts could be scaled to the transaction volume and to other proposals which would ensure appropriate treatment of small sample results. *See* Verizon CA Comments at 14 (December 18, 2000). Parties will have an opportunity to propose and discuss different treatments of the outcomes from different sample sizes.

Percentage and rate-based measures

The fundamental problem with small sample sizes for average-based parity measures is that they fail to satisfy the normality assumptions for the Modified Z-test. In contrast, percentage and rate-based measures are assessed using exact tests, which do not depend on inferences or assumptions about underlying distributions. Consequently, with these tests there is less concern with the accuracy of the alpha and beta calculations for small samples. We find no other compelling reason to aggregate or discard data, and thus, we direct that all percentage and rate-based data at the submeasure level for each CLEC be analyzed for parity regardless of sample size.[163]

Data transformations

Pacific proposes a Modified Z-test enhancement to address the data non-normality problem for average-based measures. Pacific asserts that for lognormal data distributions, transforming raw scores to their natural logs can bring the distribution close to normality, and thus satisfy the essential assumption for using a Z-test.[164] The CLECs agree to such transformations.[165] Verizon CA and ORA accept the transformation proposal in concept, but both are reluctant to use it without further research. We agree with Verizon CA and ORA so far as the record is not clear how such transformations might affect decision accuracy. However, academic sources provide guidance. For example, one text states,

> "The logarithmic transformation is particularly effective in normalizing distributions which have positive skewness. Such distributions occur… when the criterion is in terms of a time scale, i.e., number of seconds required to complete a task."[166]

---

163 I.e., no minimums are necessary. However, per our earlier discussion, maximum sample size limits are necessary for the Fisher's Exact Test because of computational limitations.

164 Pacific Opening Brief at 8 (April 28, 2000).

165 CLECs' Reply Brief at 11 (May 5, 2000).

166 Winer, *supra* at 400 (1971).

This is precisely the type of measure on which the average-based parity performance measurement is based.[167] So from a theoretical perspective, the log transformation is appropriate and reasonable. Additionally, staff has performed analyses on several qualitatively different performance results. From these analyses, staff has concluded that a log transformation (1) brings the distributions much closer to normality, and (2) provides a reasonable interpretation of skewed data. Staff's analyses of several ILEC and CLEC distributions are included as Appendix J. These analyses show the improvement when log transformations are used. In addition, they demonstrate that even in cases where the log transformation dramatically changes results from the non-transformed data, the transformed results are reasonable and appropriate treatments of the performance data.

Transformations also change the effect of outliers. For example, when an outlier exerts influence on the average result in small samples, transformations can change even the direction of the performance result from worse performance to better performance.[168] In another case, we note that the

---

[167] *See* D.099-08-020, performance measure nos. 1, 7, 14, 21, 28, 37, and 44, and staff's analysis of performance measure results frequency distributions in Appendix J.

[168] We take official notice of a lognormal transformation performed by staff on the example simulated dataset in this record. (Verizon CA Opening Brief at 2-9 and 2-13 to 2-17 (April 28, 2000).) The transformation is included in Appendix J, Attachment 6. The data represent performance measures where higher scores indicate worse performance For the raw data, the CLEC mean was worse than the ILEC mean, 9.94 and 8.29 respectively. The reverse was true for the transformed data. The CLEC mean was better than the ILEC mean, 1.81 and 2.03 respectively. The Modified Z-test score changed from the raw data Z of 1.39, to the transformed data Z of -1.89. The raw data alpha result was 0.083, whereas the transformed data alpha result was 0.97.

probabilities even for large samples where there should not be large differences change dramatically when scores are transformed.[169] While the data sets we reference may be unique examples, they raise questions that we should resolve, but are not in a position to entirely do so from the record in this proceeding to date. For the above reasons, we decline to order transformations of the data on a permanent basis unless the record is adequately developed in subsequent phases of this proceeding. Additionally, our preference is that more exact tests be used, if appropriate, which solve the small sample normality problems without transformations.

However, since we must still use the Modified Z-test, and since we must apply it to samples where normality can not be assumed, then we find that the log transformation is reasonable and appropriate, and is at least as an interim solution is necessary for application of the test to small to moderately large samples. We also find that the transformation improves normality for large samples.[170] Therefore, we shall order that log transformations be utilized for all average-based performance measures as specified in staff's analysis in Appendix J.

---

[169] We take official notice of a submeasure analysis for actual February 2000, OSS performance. With a CLEC sample size of approximately 500 and an ILEC sample size of 6,340, a Modified Z-test on raw scores produces an alpha of 0.85, whereas a Modified Z-test on transformed scores produces an alpha less than 0.0001. The difference is interpreted as follows: Raw score analysis indicates about seven to one odds that the result is due to random variation, whereas the transformed score analysis indicates there is virtually no chance that the result is due to random variation.

[170] App. J, Attachment 2.

This still leaves us with the issue of the meaning of outliers. If the impact of outliers should be minimized in our performance assessment, then the log transformations accomplish this and nothing further needs our attention. However, if outliers are meaningful in their own right, then we need to address the issue. As stated above, it is plausible that an outlier can have a disproportionate affect on competition when in the CLEC sample. Very long provisioning times could gain notoriety that could harm the reputation of a CLEC. On the other hand, outliers in the ILEC results could raise the mean and mask the fact that the ILEC is providing predominately superior service to its own customers. We believe this issue should be discussed in the incentives phase of this proceeding, and we will be open to proposals for a separate treatment of outliers in their own right. But even if parties do not propose a separate treatment of outliers or agree on their meaning, we are convinced that the log transformations provide a more appropriate Modified Z-test application. If further deliberations and negotiations of the parties do not result in adequate development of permutation testing or outlier treatment, at this point we accept the fact that log transformations may become the permanent solution.

Benchmark issues

In contrast to the parity standard for CLEC performance results with ILEC retail analogues, where there is no retail analogue, the standard is performance that allows a "meaningful opportunity to compete." In the performance measurement phase of this proceeding, the parties agreed to establish "benchmarks" which specify such performance levels.[171] Since there is no

---

[171] *See* D.99-08-020, *mimeo.* at 5-6 (August 5, 1999).

measure of an ILEC's internal performance (i.e., no retail analogue), there is no ILEC variability on which to base an expected performance parity standard. Consequently, parties negotiated measures with thresholds that would allow CLEC service access judged to allow a meaningful opportunity to compete.

The parties discussed two contentious issues regarding benchmarks. The parties discussed alleged problems of small sample sizes causing falsely missed benchmarks and random variation causing falsely missed benchmarks. Pacific proposed using adjustment tables to remedy the sample size issue and statistical testing to remedy the random variation issue.

Benchmark adjustment tables

Pacific contends that performance measures for small samples present problems in that some benchmarks would not be met even though an ILEC provided adequate service. For example, if a benchmark established that 90 percent of orders for a particular service must be complete within a certain timeframe, then for every 100 orders there could be 10 missed timeframes without failing the benchmark. Pacific points out that for small samples, one failure could drop performance below the 90 percent level. For example, if only five orders were made per CLEC, then across 20 CLECs (100 orders) there could be 10 missed timeframes (90 percent on time) and for this aggregate performance a "meaningful opportunity to compete" could be assumed by original agreement of the parties. However, at least two and at most ten CLECs in this example would have missed the benchmark. That is, if ten CLECs missed one timeframe each (for a total of 10 missed timeframes), then they each would have performance measure results of 80 percent. At least two CLECs would have to fail the performance measures (5 failures each for the total of 10 missed timeframes) even though performance was right at the benchmark.

Recognizing this problem, the CLECs have agreed to allow adjustments to the benchmark outcomes, although not to the extent desired by Pacific. Noting that benchmarks were created under the federal definition of performance allowing a "meaningful opportunity to compete,"[172] we are reluctant to allow less than the levels set by the benchmarks. To do so suggests less than a "meaningful opportunity to compete." However, in this case, because of the legitimacy of the small sample problem, and since the CLECs have agreed to some adjustments, we shall include an adjustment table in our decision model. Although the ILECs and the CLECs agree to use a benchmark adjustment table, they disagree on two aspects of such tables, sample sizes to which they will be applied and sample sizes from which they will be derived.

For the application of the adjustment tables to benchmarks results, the CLECs agree to the use of adjustment tables up to a performance result sample size of 30, and propose they be used down to a sample size of 1.[173] The ILECs propose using the tables for performance result sample sizes up to 100, down to 10 with no aggregation, and down to five with the aggregation rules they proposed for parity measures as discussed above.[174] The difference between the two proposals appears to be the type of problem they address. The CLEC table proposal appears to be addressing more closely the data "granularity"

---

172 *Id.*

173 CLECs' Reply Brief at 14-16 (May 5, 2000).

174 Pacific's Reply Brief at 4-7 (May 5, 2000); Pacific's Opening Brief at 12 (April 28, 2000); Verizon CA's Reply Brief at 11 (May 5, 2000).

problem[175] as we have described above, whereas the Pacific table proposal appears to go beyond data granularity and address broader statistical applications to benchmarks as we discuss below.

The ILECs and the CLECs also differ on the second issue, the adjustment table derivation sample size. The CLECs argue that since the table will be used on small samples, the tables should not be derived from larger samples. While they wish to limit the table's application to samples of 30, as a compromise they offer to base the table's derivation on a sample size of 100. Pacific wishes to derive the table from a sample size of 1000, but offers a derivation sample size of 400 as an alternative. Pacific states that a derivation sample size of 400 or 1000 is appropriate because the "implied performance" resulting from these derivation sample sizes is closer to the benchmark and is not unreasonably larger as would be the case with the CLEC's proposed derivation sample sizes.

While the CLECs' position is intuitively attractive in terms of the *construction* of the table, we appreciate Pacific's analysis because it assesses at least one *net effect* of the table. However, just as we are concerned with inferential statistical testing issues, we are concerned that other essential net effects have not been considered, namely the net effect that adjustment tables have in lowering the effective benchmark levels. For example, Pacific's adjustment table would allow performance to drop well below the nominal benchmarks without any failures being identified. Where the adjustment tables

---

[175] CLECs' Reply Brief at 14 (May 5, 2000).

are applied, performance could average as low as 82 percent or lower across all performance results.[176]

Additionally, we are concerned that "one size fits all" application and derivation sample size specifications may not be appropriate. For example, we note that the smallest application sample size where a whole integer failure matches the nominal 90 percent benchmark limit is 10, yet the similar smallest sample size for the nominal 99 percent benchmark is 100.[177] We find it appropriate to set different application sample sizes for different benchmark percentage levels. In the same manner, we find that a fixed derivation sample size results in varying levels of implied performance relative to the benchmark limit. For example, a derivation sample size of 400 for the nominal 90 percent benchmark results in a 92.9 percent implied performance level, which is a 29 percent movement toward perfect performance.[178] In contrast, the same derivation sample size of 400 applied to the nominal 99 percent benchmark results in a 99.68 implied performance level, which is a 68 percent movement toward perfect performance.[179] We find that the appropriate application and derivation sample sizes vary with the benchmark level.

---

[176] *See* Appendix K.

[177] One failure in 10 equals 90 percent success. One failure in 100 equals 99 percent success.

[178] *See* Pacific Reply Brief. at 5 (May 5, 2000) A 92.9 level is 30 percent of the interval between 90 and 100 percent.

[179] *Id.* A 99.68 level is 68 percent of the interval between 99 and 100 percent.

Inseparable from the problem of the granularity of the data affecting the implied performance is the affect that any adjustment will have on the established benchmarks. For example if one miss is allowed for a nominal 90 percent benchmark when applied to a sample size of five, then the benchmark percentage is effectively changed to 80 percent. Using the example of 20 CLECs with samples of five cases each as discussed above, all 20 CLECs can experience 80 percent performance without failures being identified. The overall performance for the total submeasure would be 10 percent below the nominal benchmark.

Staff has summarized the net changes to the nominal benchmarks in Appendix K. It is clear that when the adjustment tables are used, the benchmarks are substantially lowered. Recognizing these potential changes, we conclude that the implied performance level should set to address what is analogous to a Type I error without disproportionately increasing what is analogous to a Type II error. In other words, the implied performance level allowance should be higher from the nominal benchmark to a similar degree as the adjusted benchmark is effectively lowered from the nominal benchmark. With this balance in mind, we find that the application and derivation sample sizes recommended by staff in Appendix K, are more appropriate than the parties' proposals. Consequently we shall order the ILECs to use the small sample adjustment tables presented in Appendix K.

In comments to the draft decision, the CLECs object to the size of the application and derivation sample sizes stating that they are larger than necessary to address granularity. AT&T, et. al. Comments at 6 (December 18, 2000). However, we point out that because of granularity (i.e., integers) without adjustment tables the net effective percentage criterion is always higher than the nominal percentage except when the sample size is an exact multiple of the allowed missed percentage. (For example, sample sizes 10, 20, 30, 40… allow 90 percent net percentage results for the benchmark that allow 10 percent misses – the 90 percent benchmark. See Appendix K for a discussion.) We have made a judgement to address only some of that granularity, limiting our adjustment with the explicit criteria described in Appendix K.

The CLECs also object to the new tables fundamentally because they “harm CLECs by allowing more misses before finding a violation of the benchmark.” AT&T, et. al. Comments at 20 (December 18, 2000). The CLECs fail to consider that compared to both the ILEC and CLEC proposals, our application of these tables is more restrictive. Any time the CLEC industry-wide aggregate fails the benchmark, the adjustment tables are not used for CLEC-specific assessment. Our application is tailored to address conditions where actual performance result information indicates granularity most likely is a problem. (See Appendix K.)

Benchmark statistical testing

Pacific and Verizon CA also favor complete statistical testing for all benchmarks. They assert that benchmarks are subject to the same random variation problems as are parity measures. However, Pacific only acknowledges the effect of random variation on alpha and only presents remedies for alpha. We are concerned that these adjustments increase beta, and since we are at least as concerned about effects on beta, we are reluctant to make the statistical adjustments recommended by Pacific. Additionally, we interpret benchmarks to be absolute performance limits that define a “meaningful opportunity to compete.” Pacific argues that the benchmarks were created before statisticians were involved and before performance data was available, and thus the "negotiators relied on their experience in telephony and the needs of the CLECs to arrive at plausible benchmarks," and "did not fully appreciate. . . or consider. . the potential effects of random variation. . . ."[180] Yet Pacific goes on to admit that benchmarks were set recognizing that "the process in question is not completely controllable." (Id.) Pacific's speculation about what was in the minds of the negotiators is contradictory and unpersuasive. We have no confidence in basing a new statistical overlay on such speculation, as we similarly have no confidence in rejecting telephony expertise for statistical expertise.

It is clear to us that the benchmarks already allow for some random variation – no benchmark requires all services to be completed within a certain time period, and no benchmark sets a limit on the degree of any one service's outcome. For example, if the benchmark is 90% of orders completed within 4 days, and 92 percent of the actual orders were completed in 4 days or less, then Pacific is not held accountable for the random or even non-random variation of the remaining 8 percent. It would make no difference in the remedies plan whether these orders were completed within 5 or 100 days.

We are concerned that adding any additional tolerance margin to existing tolerance margins would allow two or three bites at the same apple. We prefer that if the benchmarks are not consistent with their definition of performance that will allow "a meaningful opportunity to compete," that they be adjusted directly, rather than add all the complexities and ambiguities that a new statistical overlay would create. With the inclusion of the adjustment tables we

---

[180] Pacific's Reply Brief at 4 (May 5, 2000).

specify above, we shall order that benchmarks be treated as tolerance limits. This is an issue that may be re-examined in the incentive payment phase.

Benchmark modification

Closely related to the problems that the adjustment tables and statistical tables are intended to address is the benchmark levels themselves. One possible view is that instead of using adjustment tables that the benchmarks themselves be adjusted. However, since the adjustment depends on the sample size, different benchmarks would have to be set for different sample sizes. This would be virtually the same as using adjustment tables with the current benchmarks. Consequently, we will not order a review and revision of the benchmarks at this time.

Correlation analysis

All parties agree that performance measures that are correlated because they are redundant should be treated so that multiple payments are not made for the same failure. At the same time, parties recognize that a statistical correlation alone cannot distinguish between failure redundancy and multiple instances of independent discrimination. No party wishes to implement a self-executing statistical correlation component to reduce payment for discrimination. Since our immediate concern here is for the self-executing performance remedies plan, we do not order any statistical correlation component to our decision model at this time.

We also find that parties presented correlation analysis only as an abstract concept. No implementable plans were described or proposed. If any party wishes for us to consider a correlation plan we ask that they describe a plan down to the level of detail that will allow implementation. For example, it will be important to understand what data will be analyzed, what analyses will be employed, what decision criteria will be used, and what follow-up will be used to distinguish redundancy from multiple discrimination. The plans should provide numerical examples so there is no misunderstanding about the necessary specificity of the plan.

Historical data

While our discussion here has necessarily focused on ILEC performance relative to CLEC performance at fixed time periods, ORA raises important issues about absolute performance levels. It is concerned that ILEC performance, and thus performance on behalf of the CLECs, could deteriorate over time, possibly because an ILEC's OSS systems were not constructed sufficient to handle the necessary CLEC business. Consequently, ORA is concerned that ratepayers would suffer poorer service overall, which could offset any gains that the new competitive market could provide. We agree that this is a legitimate concern, and in another phase of our review of Pacific's Section 271 application we have instituted volume testing to address this concern. However, we realize that even the best-designed test cannot anticipate all future variables. While we do not currently have anything in the record to support ordering any self-executing historical data-tracking incentives model component, we will be asking the parties to add monitoring capability to the overall plan. We shall order that at a minimum, certain performance data be monitored and analyzed for trends over time. We shall direct the parties to present proposals by the end of the trial period that would accomplish this monitoring and analysis.

Identical models for ILECs

The two ILECs, Pacific and Verizon CA, differ on an important component of our decision model. Pacific prefers to use the Modified Z-test for average-based measures whereas Verizon CA prefers to use permutation testing for these measures. We considered creating two different versions of our model to accommodate these preferences, but have decided to require the same model for both ILECs.

We have carefully analyzed all proposed model elements and have made the selections most consistent with our selection criteria. As such, our model represents the best model we could specify from the information in this record.

Additionally, since Verizon CA will in effect be a CLEC seeking access to Pacific's OSS services, and Pacific will in the same manner be a CLEC seeking access to Verizon CA's OSS services, it would not fit our criterion of fairness to allow different performance assessment methods for the two ILECs. For competition to be optimal, the playing field must be as level as possible. The two ILECs must be held to the same standard. For example, it is likely that for some average-based measures, given the same results, the permutation test would show the results as a "pass" while the Modified Z-test would show the same result as a "failure." For the above reasons, we order the same decision model for both ILECs.

Payment retroactivity

Verizon CA asks that the Commission hold any performance remedies plan incentive payments in an escrow account until the end of the trial period. However, since we expect that Pacific will be making its Section 271 application on the basis of the trial period having a self-executing performance remedies plan, we do not wish to allow retroactive adjustments. To do so would in essence nullify the self-executing nature of the plan. In other words, a self-executing plan is one that will trigger incentive payments without any new decisions; the decision model automatically makes decisions. If retroactive changes are made after new consideration, debate, and decisions, then the plan is not truly self-executing. We are also concerned that allowing retroactive payment alteration will make the already difficult decision model development task more cumbersome.

Some "calibration" with actual data will be helpful in assessing our decision model and its effects on the overall plan, and we will order a calibration period to occur simultaneously with the incentive payment setting phase of this proceeding before the trial period begins. We are concerned that retroactively allowing payment amounts to be adjusted at the end of the trial period will cause the parties' positions regarding the appropriateness of the decision model to be too influenced by their own corporate outcomes, relative to being influenced by the criteria we have described herein. For the above reasons, the trial incentive payments shall be made consistent with the self-executing function of the plan to be determined before the trial period begins. Incentive payment amounts shall not be altered retroactively unless we specifically provide for such alteration in the final plan.[181] In comments on the draft decision, Verizon raises legal questions that we intend to resolve before a final plan is adopted. Verizon Comments at 5, 16-19.

---

[181] Our discussion and decision on retroactivity does not address the issue of the correction of mistakes in the data or calculations necessary to arrive at incentive payments. This correction issue should be resolved in the incentives phase of these proceedings.

Other issues

Z-statistic negative/positive interpretation

The Modified Z-test statistic becomes a negative or positive value depending on whether the average CLEC performance measurement result (mean) is larger or smaller than the ILEC result (mean), and depending upon whether the CLEC mean is subtracted from the ILEC mean or vice-versa.[182] We note that potential[183] non-parity performance is represented by a negative Z-statistic in both the New York remedies plan and the Louisiana proposed remedies plan and by a positive Z-statistic in the Texas plan. While there would be some merit in constructing our decision model to be consistent with other states, given the already established inconsistency, we must base our decision on some other criterion. We prefer the convention that is most likely understood by those with little statistical sophistication. Because the typical connotations of the words “negative,” “discrimination,” and “failure,” are similar, and the

[182] For the sake of this illustration, assume the average time taken for Pacific to provision a hypothetical service for its own customers is 7 days and the average time taken for Pacific to provision service for a CLEC customer is 14 days. In this case, a longer time is worse performance and could create a barrier to competition. If the ILEC mean is subtracted from the CLEC mean (14 – 7 = +7), then a positive Z-test statistic represents a potential non-parity condition. But if the CLEC mean is subtracted from the ILEC mean (7 – 14 = -7), then a negative Z-test statistic represents a potential non-parity condition. This would be reversed for measures where a larger number represents better performance. For consistency in the interpretation of the Z-statistic, the order of the means (i.e., which mean is the subtrahend) must be reversed for situations where larger numbers represent worse performance compared to situations where larger numbers represent better performance.

[183] We use the term “potential” here because non-parity identification will also depend on the magnitude of the Z-statistic (i.e., it must be either a larger positive value than a positive critical value or a larger negative value than a negative critical value).

connotation of “positive” is opposite from these other words, we prefer the Z-test be implemented with a negative Z-value representing potential discrimination. Reading “negative” values to represent negative outcomes is intuitively understandable whereas the reverse is not. Therefore, we shall order our decision model constructed so that negative Z-values represent potential discrimination.

Performance Measure 42

In comments to the draft decision, Pacific pointed out that Performance Measure 42 was unique, and that proposed statistical tests could not be appropriately applied. Pacific Comments at 3 (December 18, 2000). Pacific proposed that for the parity submeasures within Measure 42, “the ILEC percentage minus the CLEC percentage should not exceed 0.05 percentage points. (Reflected in proportions, this difference would be 0.0005).”[184] ORA agrees that Pacific’s proposal is appropriate. ORA Reply Comments at 5 (December 22, 2000). As other parties are silent regarding Pacific’s proposal, we assume no objections. As Pacific’s proposal seems reasonable and has either explicit or implicit concurrence of other parties, we shall include it as part of the decision model we adopt today.

---

[184] For example, for “systems available 500 hours during a month, this difference translates into a total discrepancy of 15 minutes.” Pacific Comments at 4 (December 18, 2000).

Parity performance measures without sufficient ILEC data

Parity comparisons cannot be made without ILEC performance data. Since there may be insufficient ILEC activity in some months for some measures, we need to specify alternative retail analogues. Tests that require standard deviation calculation require at least two observations and exact tests require at least one observation. Pacific proposes that the prior six months of ILEC data be aggregated (to the extent that such data exist) and used in place of the data-deficient month, and if the aggregate does not produce sufficient ILEC data, the submeasure not be evaluated for the month. Pacific Comments at 19 (December 18, 2000). The CLECs agree with the exception that they wish to use the prior three months CLEC data as a surrogate analogue instead of failing to evaluate the performance results. AT&T, et. al. Reply Comments at 3 (December 22, 2000). We agree with Pacific's proposal. Using historical CLEC data may confound discriminatory behavior with seasonal fluctuations. If there is no retail analogue for six months, parties should consider creating a benchmark to assess performance.

Interim and permanent models

As recommended by the ACR, the model we now adopt is an interim model that will generate incentive payments once we have added the incentive components in the next phase of this proceeding. After six month's experience with the model we will review its performance and adjust any component that we find needs changing. Implementing this model as a fully functioning and self-executing performance remedies plan will allow Pacific to file its section 271 application for entry into the in-region interLATA long distance market. At the same time, this trial period will allow actual experience to guide future refinements. While any party can at any time petition us to change the model, we will remove that burden of persuasion by scheduling this review and adjustment opportunity. As discussed in detail above, there are many unresolved issues regarding what would be the best and most appropriate model. We find that we cannot resolve all these issues. Yet at the same time, we conclude that we can proceed with a fully implementable model while gaining the experience necessary for future development of a permanent model.

Comments on Draft Decision

The draft decision of ALJ Jacqueline A. Reed in this matter was mailed to parties in accordance with Pub. Util. Code § 311(g)(1) and Rule 77.7 of the Rules of Practice and Procedure. Comments were filed on December 18, 2000, and reply comments were filed on December 22, 2000. We have taken the comments into account, as appropriate, in finalizing this decision. As this is an interim decision, there will be an opportunity for us to consider and implement modifications before a final decision is adopted.

Findings of Fact

1. The cornerstone of any performance incentive structure is how parity is defined, since it is on those occasions when an ILEC is out of parity that incentive payments will be made.

2. This Commission's definition of parity incorporates the objectives of the TA96 and the FCC.

3. It will be helpful to rely on statistical testing and benchmarks to infer whether or not parity has been achieved.

4. In late fall 1999, the existent ILEC models and the CLECs' model were distinct and irreconcilable.

5. The parties revealed considerable misunderstanding and confusion about the two sets of respective model assumptions and calculations.

6. The outcomes of the two models were highly discrepant because both approaches were trying simultaneously to design and implement the total model (both the performance assessment model elements and the incentive plan elements) without the benefit of an implementation and data calibration structure.

7. It is unlikely that either model could be implemented as designed.

8. During the February 1999 technical workshop, each proposed plan produced dramatically different payments due to different input assumptions.

9. There is a need to have one common interim model framework of analyses for review and discussion in order to implement the performance remedies plan.

10. To achieve a common model framework, the performance assessment model elements and the incentive plan elements need to be separated.

11. Since the task of accurately assessing the state of competitive conditions must be self-executing, the decision model must be able to automatically identify performance result levels that reveal competition barriers and that will trigger incentive payments.

12. There are two fundamental categories of performance measures that must be assessed to determine the existence of competitive conditions: “parity” and “benchmark” measures.

13. In identifying parity or non-parity, accurate remedies-plan decision-making involves more than accurately calculating average ILEC and CLEC

performance and identifying non-parity if ILEC service to CLEC customers is significantly worse than ILEC service to ILEC customers.

14. Given that there is variability in ILEC performance in providing retail services to its own customers, a measurement showing inferior service to CLEC customers could be due either to this variability, or actual discrimination, or both.

15. Statistical testing allows estimation of decision accuracy, or in other words, calculation of the decision error probabilities.

16. These probabilities can then assist decision-making by quantifying the different error probabilities and comparing them to standards of confidence that the Commission wishes to apply.

17. Using measures of performance averages and variability, statistical analysis provides estimates of: (1) the probability that a result of a certain magnitude would be detected when it exists (test power and corresponding error beta) and (2) the probability that the result is due to random variation when in fact there are no differences (confidence level and corresponding error alpha).

18. Benchmarks have been constructed as tolerance limits.

19. The issues for statistical analysis accuracy of benchmarks are not the same as those for parity measures.

20. None of the presented models for parity assessment are acceptable in their entirety.

21. Four types of measurements have been developed for monitoring ILEC performance: averages, percentages, indexed and rates.

22. Each measurement type requires a different statistical test or a variant of the same test.

23. All parties have agreed that a one-tailed statistical test should be used.

24. In response to the CLECs' concerns that ILEC discrimination could increase the CLEC variance, and thus make it more difficult to detect any discrimination, all parties agreed to use a Modified Z-test instead of the standard Z-test.

25. According to the statistical literature, requiring normally distributed data in the use of any Z-test may be only partially correct.

26. The Central Limit Theorem states that for sufficiently large samples, non-normality in the data does not affect the test.

27. The permutation test has the potential for being a more accurate test that can handle small samples.

28. The Z-test relies on the resulting sampling distributions being approximately normal.

29. The few data examples we have available to us comparing permutation and Z-tests show the expected divergence for small samples, but not the expected convergence for larger samples, contrary to the theoretical expectation that the results should be the same for large sample sizes.

30. The results of the few available data examples raise doubts that the record is sufficiently developed to allow the Commission to confidently select the permutation test as a superior test for average-based measures.

31. In the interim, the Z-test is the most developed and accepted alternative to permutation testing.

32. The advantage of exact tests for the Commission's statistical model is two-fold: (1) calculations are made directly from the raw data, and (2) the exact tests have the potential to produce more accurate results for small samples.

33. Unlike for average-based permutation applications, outliers cannot affect the result of the Fisher Exact test, as the data consist only of "cell counts."

34. Additionally, unlike for average-based permutation applications, the results from the percentage-based Modified Z-test and the results from the Fisher's Exact Test converge towards equality as theoretically expected.

35. The Fisher's Exact Test generates computationally difficult numbers that unnecessarily drain computer resources for no benefit in accuracy for large samples.

36. The Fisher's Exact Test is appropriate and can be calculated up to a limit of 1000 CLEC performance "hits" or "misses," and the Modified Z-test for proportions is appropriate for performance results above this limit.

37. Like the percentage-based Fisher's Exact test applications, and unlike for average-based permutation applications, the results from the rate-based Modified Z-test and the results from the binomial exact test converge towards equality as theoretically expected.

38. Balancing alpha and beta to be equal only ensures that the most accurate decision is made, not what the relative consequences of those decisions will be.

39. The record is relatively silent on the actual beta values that various critical alpha levels might produce.

40. The record is relatively silent on the appropriate test power or beta error level.

41. The record is incomplete regarding what performance level deltas would be, because no party has submitted any proposal containing a comprehensive set of specific deltas.

42. A fixed alpha is not an adequate long-term solution.

43. Test power is very low for the small samples that represent the majority of the performance measure results.

44. Fixed alphas that provide better test power for small samples result in unnecessarily high test power for large samples.

45. A larger alpha level of 0.10, instead of the 0.05 level, enhances decision accuracy and avoids uncorrectable decision-making errors while still addressing correctable errors in the next phase of this proceeding.

46. A smaller alpha level than 0.15 is reasonable because of concerns about the effect on large-sample results.

47. An 80% confidence level (0.20 alpha) in the model for conditional failure identifications is warranted because of the high beta error still remaining when using the 0.10 alpha level.

48. Both record efforts to establish “material” thresholds have merit.

49. The “material difference” standard has merit and the potential to improve the decision model we specify.

50. Minimum sample size requirements vary depending upon the type of statistical test used.

51. Harmful ILEC performance in small new or innovative market niches, or harmful ILEC performance to smaller CLECs, could be masked by relying on assessments of larger market samples or larger CLEC samples when the results for CLECs are aggregated.

52. It is important to examine performance at the smaller market and smaller CLEC levels.

53. There are unresolved issues regarding minimum sample size and sample aggregation rules, and the rules for incentive payments are integrated with the aggregation rules.

54. Minimum sample size rules result in some data being discarded.

55. Our small sample aggregation rules avoid discarding any data and increase sample sizes for the very smallest samples with minimal impact on the actual results.

56. The previously proposed sample size rules are complicated and fall short of our goal of simplicity.

57. The fundamental problem with small sample sizes for parity measures is that they fail to satisfy the normality assumptions for the Modified Z or *t*-test.

58. Statistical texts indicate that the *t*-distribution is more appropriate than the Z-distribution for tests between two sample means, especially for small samples.

59. Using the t-distribution table would adjust for decreasing sample size.

60. Percentage and rate-based measures are assessed using exact tests, which do not depend on inferences or assumptions about underlying distributions.

61. A log transformation (1) brings the distributions much closer to normality, and (2) provides a reasonable interpretation of skewed data.

62. ILEC distribution normality is improved when log transformations are used.

63. Log transformations also change the effect of outliers.

64. Log transformation improves normality for large samples.

65. Log transformations provide a more appropriate Modified *t*-test application than an application using data that is not transformed.

66. Although the ILECs and the CLECs agree to use a benchmark adjustment table, they disagree on two aspects of such tables, sample sizes to which they will be applied and sample sizes from which they will be derived.

67. A fixed derivation sample size results in varying levels of increased implied performance relative to the benchmark limit.

68. The appropriate application and derivation sample sizes vary with the benchmark level.

69. When the adjustment tables are used, the benchmarks are substantially lowered.

70. The application and derivation sample sizes recommended by staff in Appendix K, are more appropriate than the parties' proposals.

71. Benchmarks are absolute performance limits that define a "meaningful opportunity to compete."

72. Benchmarks already allow for some random variation – no benchmark requires all services to be completed within a certain time period, and no benchmark sets an upper limit on any one service's outcome.

73. Performance measures that are correlated because they are redundant should be treated so that multiple payments are not made for the same failure.

74. No party wishes to implement a self-executing statistical correlation component to reduce payment for discrimination.

75. Parties presented correlation analysis only as an abstract concept; no implementable plans were described or proposed.

76. Allowing retroactive adjustments would nullify the self-executing nature of the performance remedies plan.

77. Reading "negative" values to represent negative outcomes is intuitively understandable whereas the reverse is not.

78. A special index must be created for performance measure 42 since the proposed parity statistical tests cannot be appropriately applied.

79. Parity comparisons cannot be made without ILEC performance data and alternative retail analogues must be created for months where there is insufficient ILEC data.

80. Tests that require standard deviation calculation require at least two observations and exact tests require at least one observation.

81. Using the prior six months of aggregated ILEC data be aggregated (to the extent that such data exist) is an appropriate alternative retail analogue.

82. Using historical CLEC data as a surrogate for a retail analogue may confound discriminatory behavior with seasonal fluctuations.

83. The present fully implementable model is an interim one that will generate incentive payments once we have added the incentive components in the next phase of this proceeding.

Conclusions of Law

1. Parity means that the ILEC is providing services to the CLECs in substantially the same period of time and manner (including quality) as it is providing to itself.

2. This Commission endeavors to ensure that the CLECs have OSS access that is at least equal to the ILECs' own access.

3. One interim performance remedies plan model and set of explicit assumptions would allow common quantitative analyses to be performed and estimates to be developed.

4. A single model approach would allow the Commission to make informed and fair policy decisions about the performance remedies plan.

5. A single model approach focuses on the goal of parity service by the ILECs, economic incentives paid by the ILECs, and/or a change in ILECs' operations support to the CLECs.

6. A single interim model and a single set of explicit assumptions should allow calibration of economic outcomes both before and after a six-month pilot test period using actual empirical data.

7. The interim pilot test period will assist the Commission in determining the appropriate levels of long-term economic incentives.

8. Long-term incentive impacts can be calibrated in relation to one model, one common set of assumptions, and actual test period empirical data.

9. Statistical testing should be used to assess the balance between finding and preventing actual barriers, and avoiding the identification of barriers when they do not exist, thus enabling greater decision quality and attainment of legislative goals.

10. A new "hybrid" of elements from each of the different models presented in this proceeding constitutes the most appropriate performance remedies statistical model.

11. Consistent with academic texts and with the FCC's view of the appropriate statistical application regarding the requirements of the Act, a one-tailed test is appropriate for situations where there is only interest in outcomes in one direction, in this case where the CLEC performance results are worse than the ILEC results.

12. The selection of the appropriate test for small samples should be based on the relative accuracy of the different tests.

13. It is reasonable for our sample aggregation rules to act as an interim solution and a "floor" for sample sizes.

14. Evidence in this proceeding is compelling that normality cannot be assumed for small samples since measures of time-delay are commonly skewed – the distribution is "bunched up" for shorter delays, and tapers off slowly for longer delays.

15. Until the Commission can determine which test is the more appropriate treatment of the data, including underlying issues such as "production output"

versus “larger process population sampling” and more specific issues regarding outlier treatment, it is not reasonable to either approve or order use of the permutation test.

16. There is a need to better understand what the appropriate sample sizes are for using the permutation test versus the Modified Z or *t*-test.

17. Since there are unresolved questions surrounding the potential of the permutation test, the active interested parties in this proceeding should collaboratively conduct or fund a research inquiry to answer these unresolved questions.

18. In the case of the percentage-based performance results data, the Fisher’s Exact test is appropriate.

19. The Fisher's Exact test should be used for percentage-based performance results because it provides accurate decision error probabilities, is consistent with theoretical assumptions, solves the Z-test application problems.

20. The binomial exact test should be used for rate-based performance results because it provides accurate decision error probabilities, is consistent with theoretical exceptions, solves the Z-test application problems, is preferred by most parties.

21. The question of relative risk is more appropriately addressed in this proceeding’s next phase, which will establish the “consequences” for the performance decisions made in the present phase.

22. To remedy the lack of critical record information, it is reasonable to direct the ILECs to calculate both alpha and beta values whenever a statistical test is applied.

23. As a general policy statement, it is reasonable to assume that a Type II error is at least as important as a Type I error. Apparent discrepancies can be adjusted in the incentive payment phase.

24. It is reasonable that the problems of insufficient test power for small samples (large beta) and “too much” test power for large samples can be better resolved through even approximate alpha / beta balancing techniques.

25. A fixed alpha critical value should only be used in the model as an interim decision-criterion solution.

26. The 90% confidence level (0.10 alpha, or 10% significance level) should be adopted in the statistical model to control the Type I error and to reduce the Type II error to more acceptable levels for the preponderance of the performance results.

27. Pending establishment of applicable conditions, the 80% confidence level (0.20 alpha) should be adopted in the statistical model for conditional failure identifications because of the low power of these tests.

28. The parties should be directed to devise and propose specific conditional failure identifications in the next phase of this proceeding.

29. One goal of the performance remedies plan is to assess each CLEC’s performance results for each submeasure.

30. The smaller market and smaller CLEC levels may be critical for entry and innovation, which in turn are critical to a healthy competitive telecommunications infrastructure.

31. Consistent with the academic justification of the Modified Z-test, the test statistic should be compared to the *t*-distribution.

32. The small sample aggregation rules we have designed should be easily understood with the results easily reproduced.

33. To assess performance subject to the performance remedies plan, statistical analysis and decision rules should be applied to all data, including sample sizes as small as one case, after our small sample aggregation rules are applied.

34. How payments will be triggered or allocated under the aggregation rules should be addressed in the upcoming incentives phase.

35. All percentage and rate-based data at the submeasure level for each CLEC should be analyzed for parity regardless of small sample sizes since exact tests are accurate for all sample sizes.

36. Staff's analyses of several ILEC and CLEC distributions demonstrate that even in cases where the log transformation dramatically changes results from the non-transformed data, the transformed results are reasonable and appropriate treatments of the performance data.

37. Log transformations of the data should not be ordered on a permanent basis until the record is adequately developed in subsequent phases of this proceeding.

38. More exact tests should be used in addressing small sample size issues, if subsequent research shows them to be appropriate.

39. The log transformation is reasonable and appropriate, and is necessary at least as an interim solution for application of the Modified Z-test to small to moderately large samples.

40. Log transformations should be utilized for all average-based performance measures as specified in Appendix J.

41. The meaning of outliers should be discussed in the incentives phase of this proceeding.

42. Because of the legitimacy of the benchmark small sample problem, and since the CLECs have agreed to some adjustments, a benchmark small sample adjustment table should be ordered as part of the decision model.

43. It is appropriate to set different application sample sizes for different benchmark percentage levels.

44. The implied performance level should be set to address what is analogous to a Type I error without disproportionately increasing what is analogous to a Type II error.

45. The ILECs should use the small sample adjustment tables presented in Appendix K.

46. If any benchmark is inconsistent with the performance definition "a meaningful opportunity to compete," it should be adjusted directly rather than add all the complexities and ambiguities that a new statistical overlay would create.

47. Benchmarks should be treated as tolerance limits; however, the issue may be re-examined in the incentive payment phase.

48. A review and revision of the benchmarks should not be ordered at this time because it could be more cumbersome than using adjustment tables with the current benchmarks, and establishing benchmarks is the subject of a different proceeding.

49. Since parties recognize that a statistical correlation alone cannot distinguish between failure redundancy and multiple instances of independent discrimination, we should not order any statistical correlation component to our self-executing performance remedies plan model.

50. Any party seeking to have a correlation plan considered in the next phase of this proceeding should describe the plan down to the level of detail that will

allow implementation. Parties should provide numerical examples so there is no misunderstanding about the necessary specificity of the plan.

51. The parties should present proposals by the end of the trial period that would put into effect the monitoring and analysis of certain performance data for trends over time.

52. The same performance remedies model should be applied to both Pacific and Verizon CA in the interest of fairness.

53. Since some "calibration" with actual data will be helpful in assessing our decision model and its effects on the overall plan, a calibration period should be ordered to occur simultaneously with the incentive payment setting phase of this proceeding before the trial period begins.

54. Allowing retroactive payment alteration will make the already difficult decision model development task more cumbersome.

55. Incentive payment amounts should not be altered retroactively.

56. Following a six-month trial period, to be specified in the incentive payment phase of this proceeding, the performance of the remedies plan model should be reviewed and any component determined to need changing should be adjusted.

57. A fully implementable interim model should be utilized while gaining the experience necessary for future development of a permanent model.

58. This decision should become effective immediately so that the calibration process can begin and the incentive payment phase may proceed.

INTERIM ORDER

IT IS ORDERED that:

1. A performance remedies plan decision model, which identifies performance failures and non-failures, as specified in Appendix C incorporated by reference herein, shall be adopted for Pacific Bell (Pacific) and Verizon California Inc. (Verizon CA).

2. The performance remedies plan, comprised of the decision model adopted herein and an incentive payment component that will be determined in the next phase of this proceeding, shall be implemented for a trial period of six months.

3. Pacific and Verizon CA shall use the Modified *t*-test for average-based parity performance measures.

4. Log transformations shall be utilized for all average-based performance measures as specified in Appendix J.

5. Pacific, Verizon CA and the active interested competitive local exchange carriers (CLECs) in Rulemaking 97-10-016/Investigation 97-10-017 shall collectively conduct or fund a research inquiry into whether the permutation test or the Modified *t*-test is the more appropriate treatment of the data, including but not limited to underlying issues such as “production output” versus “larger process population sampling” and more specific issues regarding outlier treatment. The inquiry shall adopt a collaborative research approach so that all interested parties can collectively influence the research proposal.

6. The Fisher’s Exact test shall be used for all percentage-based parity results except for those that cannot be computed because of large numbers. Results where the CLEC numerator exceeds 1000 shall be calculated with the Modified Z-test for proportions.

7. The binomial exact test shall be used for all rate-based tests.

8. The performance remedies plan model shall be constructed so that negative Z and *t*-values represent potential discrimination.

9. Pacific and Verizon CA shall calculate and report both Type I (alpha) and Type II (beta) error values whenever a statistical test is applied.

10. The parties shall collaboratively develop and implement an alpha/beta balancing procedure for the statistical model adopted herein and detailed in Appendix G no later than the end of the trial period, unless the parties reach agreement and jointly move to implement the components sooner.

11. If the parties are unable to agree on an alpha/beta balancing decision component for the model by the end of the trial period, the parties shall submit their individual models for Commission review and decision as directed by the assigned Commissioner and/or assigned Administrative Law Judge.

12. Until an alpha/beta balanced criterion is established, fixed alpha critical values shall be adopted for the interim.

13. A 90% confidence level (0.10 alpha, or 10% significance level) shall be adopted as the interim fixed critical value in the statistical model for failure identifications.

14. For the possible implementation of an 80% confidence level (0.20 alpha), the parties shall devise and propose specific conditional failure identifications for our consideration in the next phase of this proceeding.

15. Except for rare submeasures identified in Appendix H, Attachment 1, the following small sample aggregation rules shall be used for average-based parity performance measures: (1) For each submeasure, all samples with one to four cases shall be aggregated with each other; and (2) statistical analyses and decision rules shall be applied to determine performance subject to the

performance remedies plan for all samples after the aggregation in step (1), regardless of sample size.

16. Rare submeasures identified in Appendix H, Attachment 1, shall be analyzed without aggregation and regardless of sample size.

17. How payments will be triggered or allocated under the aggregation rules shall be addressed in the upcoming incentives phase.

18. All percentage and rate-based data at the submeasure level for each CLEC shall be analyzed for parity without aggregation and regardless of sample size.

19. Pacific and Verizon CA shall use the small sample adjustment tables presented in Appendix K.

20. Benchmarks shall be treated as tolerance limits; however, the issue may be re-examined in the incentive payment phase.

21. Pacific, Verizon CA and any interested parties shall present proposals by the end of the trial period that would put into effect the monitoring and analysis of certain performance data for trends over time.

22. The same performance remedies model shall be applied to Pacific and Verizon CA.

23. A calibration period shall occur simultaneously with the incentive payment setting phase of this proceeding before the trial period begins.

24. Following a six-month trial period, to be specified in the incentive payment phase of this proceeding, we shall review the performance of the remedies plan model and adjust any component that we determine needs changing.

This order is effective today.

Dated January 18, 2001, at San Francisco, California.

LORETTA M. LYNCH
President
HENRY M. DUQUE
CARL W. WOOD
Commissioners

Commissioner Richard A. Bilas, being necessarily absent, did not participate.

Appendix A

ACR Questions

Assigned Commissioner's Ruling Questions

Tests for Determining Compliance with Parity

1. A standardized Z-test is proposed for purposes of determining compliance with parity. Explain why this standard textbook statistical test cannot serve as a measurement tool at least for the duration of the six-month trial pilot test period? Keep in mind that the incentive phase of the model can calibrate for measurement outcomes through various incentive plan structures and amounts.
2. Benchmark measures without any statistical tests are proposed for purposes of determining a performance failure. Explain why this simple approach cannot serve as a measurement tool at least for the duration of the six-month trial pilot test period? Keep in mind that the incentive phase of the model can incorporate information on underlying data values and distributions.

Minimum Sample Sizes

1. A minimum sample size of thirty, aggregated in up to three-month time periods, is proposed. Explain why this standard textbook statistical proposal cannot serve as a minimum sample size rule at least for the duration of the six-month trial test period? Keep in mind that the test would still be performed using whatever sample size is achieved at the end of three months.

Alpha Levels/Critical Values

Ten percent Type I alpha level for parity tests is proposed. Explain why this standard textbook statistical proposal cannot serve as an alpha level/critical value rule at least for the duration of the six-month trial pilot test period? Again, keep in mind that the penalty phase of the plan can calibrate the size of the payments as a function of the critical values.

Appendix B

References

Bartz, A. (1988). *Basic statistical concepts*. 3rd ed. New York: Macmillan.

Bickel, P. & Doksum, K. (1977). *Mathematical statistics: Basic ideas and selected topics*. San Francisco: Holden-Day.

Brownie, C., Boos, D., & Hughes-Oliver, J. (1990). Modifying the t and ANOVA F tests when treatment is expected to increase variability relative to controls. *Biometrics, 46*, 259-266.

Brubaker, K. & McCuen, R. (1990) Level of significance selection in engineering analysis. *Journal of Professional Issues in Engineering, 116,* 375-387.

Das, C. (1994). Decision making by classical test procedures using an optimal level of significance. *European Journal of Operational Research*, *73*, 76-84.

Gold, D. (1969). Statistical tests and substantive significance. *The American Sociologist*, *4*, 42 – 46.

Good, P. (2000). *Permutation tests: A practical guide to resampling methods for testing hypotheses*. 2nd Ed. New York: Springer Verlag.

Hays, W. (1994). *Statistics*. 5th ed. Fort Worth: Harcourt Brace.

Hubbard, R.; Parsa, R.; Luthy, M. (1997). The spread of statistical significance testing in psychology: The case of the Journal of Applied Psychology, 1917-1994. *Theory & Psychology*, *7*, 545-554.

Hunter, J. (1997). Needed: A ban on the significance test. *Psychological Science*, *8*, 3-7.

Johnstone, D. & Lindley, D. (1995). Bayesian inference given data "significant at α": Tests of point hypothesis. *Theory & Decision*, *38*, 51 – 60.

Khazanie, R. (1997). *Statistics in a world of applications.* 4th ed. Harper Collins.

McNemar, Q. (1962). *Psychological statistics*. New York: John Wiley & Sons.

Raiffa, H. (1970). *Decision analysis*. Reading, Mass.: Addison-Wesley

Sheskin, D. (1997). *Handbook of parametric and nonparametric statistical procedures.* Boca Raton: CRC Press.

Skipper, J., Guenther, A., & Nass, G. (1970). The sacredness of .05: A note concerning the uses of statistical levels of significance in social science. *The American Sociologist*, *2*, 16-18.

Verma, R. & Goodale, J. (1995) Statistical power in operations management research. *Journal of Operations Management*, *13*, 139-152.

Welsh, A.H. (1996). *Aspects of statistical inference*. New York: Wiley & Sons.

Winer, B.J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill

Appendix C

Decision Model

I. Parity measures

A. Statistical Tests

All statistical tests will be one-tailed tests.

1. Average-based Parity Measures

The Modified *t*-test will be used for all average-based parity measures as specified in:

Brownie, C., Boos, D., & Hughes-Oliver, J. (1990). Modifying the t and ANOVA F tests when treatment is expected to increase variability relative to controls. *Biometrics*, *46*, 259-266.

The Modified *t*-test for the difference in means (averages) between the ILEC and the CLEC populations is:

$$t = (M_i - M_c) / [S_i * \text{sqrt}(1 / N_c + 1 / N_i)]$$

Where:
- M_c = the CLEC mean result
- M_i = the ILEC mean result
- S_i = the standard deviation of the results for the ILEC
- N_c = the CLEC sample size
- N_i = the ILEC sample size
- sqrt = square root

For measures of time intervals, except for data where “zeros” are not possible, the raw score distribution will be normalized by taking the natural log of each score after a constant of 0.4 of the smallest unit of measurement is added to each score. For example, if the smallest unit of measurement is an integer, then the added constant would be 0.4:

$$x_{tran} = \ln(x + 0.4)$$

Similarly, if the smallest unit of measurement is 0.01, then the added constant would be 0.004:

$$x_{tran} = \ln(x + 0.004)$$

Results that are not measures of time intervals (e.g., Measure 34) will not be transformed.

The Modified *t*-test calculation for average parity measures will be structured so that a negative sign indicates "worst" performance. Specifically, when a lower value represents better performance, such as time to provision a service, the CLEC mean will be subtracted from the ILEC mean. Different performance measures may require reversing the means in the equation to have a negative sign indicate poorer performance.

The *t*-statistic will be converted to an α (Type I error) probability using a *t*-distribution table or calculation. Degrees of freedom (*df*) will be based only on the ILEC sample size consistent with Brownie, et al. If the obtained α value is less than the critical α value, then the result will be deemed not in parity.

2. Proportion Parity Measures

Except for performance results that have numbers too large to calculate with the exact test, the Fisher's Exact Test will be used for all percentage or proportion parity measures as specified in:

Sheskin, D. (1997). *Handbook of parametric and nonparametric statistical procedures.* Boca Raton: CRC Press, pp. 221-225.

If the obtained α value is less than the critical α value, then the result will be deemed out-of-parity.

Performance results that are too large to calculate with the Fisher's exact test are those measures that exceed the following values:

1. For percentage-based measures where low values signal good service, Fisher's Exact Test shall be applied to all problems for which the CLEC numerator is less than 1000 "hits." The Z-test shall be applied to larger results.

2. For percentage-based measures where high values signal good service, the analysis is the same but is applied to the "misses" as opposed to the "hits." The Fisher's Exact Test shall be applied whenever the denominator minus the

numerator is less than 1000 for the CLEC result. The Z-test shall be applied to larger results.

Such results will be calculated using the Modified Z-test for proportions as follows:

$$Z = (P_i - P_c) / \mathrm{sqrt}[P_i(1-P_i)*(1/N_c + 1/N_i)]$$

Where:

P_c = the CLEC proportion
P_i = the ILEC proportion
N_c = the CLEC sample size
N_i = the ILEC sample size
sqrt = square root

The Modified Z-test calculation for proportion parity measures will be structured so that a negative sign indicates "worst" performance. Specifically, when a higher value represents better performance, such as percent on-time tasks, the ILEC proportion will be subtracted from the CLEC proportion. Different performance measures may require reversing the means in the equation to have a negative sign indicate poorer performance.

The Z-statistic will be converted to an α (Type I error) probability using a Z-distribution table or calculation. If the obtained α value is less than the critical α value, then the result will be deemed not in parity.

3. Rate-based Parity Measures

The Binomial Exact Test will be used for all rate parity measures. The Binomial Exact Test is specified in GTECs Exhibit C, Section 3, "Permutation Test for Rates", Equations 3.1 and 3.2 (Deliverable #7, Facilitated Work Group, April 2000).

4. Indexed-based Parity Measures

Measure 42 provides an index of parity performance that will be assessed by comparing ILEC and CLEC performance as follows:

Non-parity will be identified when the ILEC percentage minus the CLEC percentage exceeds 0.05 percentage points.

B. Critical Alpha Level for Parity Tests

The Type I error probabilities (alphas, α) obtained from the parity statistical tests will be compared to a critical alpha value of 0.10.

A performance result with α equal to or less than 0.10 will be deemed a performance failure with no additional conditions.

A performance result with α equal to or less than 0.20 and greater than 0.10 will be deemed a conditional failure. Additional conditions to determine failures will be specified in the final remedies plan.

C. Sample Sizes and Aggregation Rules

Statistical tests will be applied to the monthly performance results specified in D.99-08-020.

1. Average-based measures

For average-based performance results the following aggregation rules will be used:

(1) For each submeasure, the performance results for all samples with one to four cases will be aggregated with each other to form a single performance result.

(2) Statistical analyses and decision rules will be applied to determine performance subject to the performance remedies plan for all samples after the aggregation in step (1), regardless of sample size. For example, if samples with as few as one case remain after the aggregation, statistical analysis and decision rules will be applied to determine performance subject to the performance remedies plan to these samples, just as they are for larger samples.

2. Proportion and rate-based measures

All samples will be analyzed as they are reported without aggregation.

D. Measures without Retail Analogues.

In months where there are no retail analogue performance data, the prior six months of ILEC data be aggregated (to the extent that such data exist) and used in place of the data-deficient month. If the aggregate does not produce sufficient ILEC data, the submeasure not be evaluated for the month.

II. Benchmark Measures

For large samples, the actual performance will be compared to the benchmark nominal percentage according to the percentage set in the Joint Partial Settlement Agreement approved by the Commission. For small samples, maximum permitted “misses” shall be determined by small sample adjustment tables. Small samples are defined as follows:

90 percent benchmarks - 50 cases or less
95 percent benchmarks - 100 cases or less
99 percent benchmarks - 500 cases or less

Adjustment tables:

90% Benchmark		95% Benchmark		99% Benchmark	
Sample size	Maximum permitted misses	Sample size	Maximum permitted misses	Sample size	Maximum permitted misses
1	0	1 to 3	0	1 to 19	0
2 to 9	1	4 to 19	1	20 to 97	1
10 to 20	2	20 to 40	2	98 to 202	2
21 to 31	3	41 to 63	3	203 to 319	3
32 to 44	4	64 to 88	4	320 to 445	4
45 to 50	5	89 to 100	5	446 to 500	5

The small sample adjustment tables shall be used in the following steps:

1. The number of performance "misses" for the CLEC industry-wide aggregate for each remedy plan benchmark submeasure will be compared to the number of permitted misses for all sample sizes covered by the related adjustment table. Industry aggregate performance will be identified as passing if the number of actual misses is less than or equal to the number of permitted misses, and identified as failing if otherwise.

2. For CLEC industry-wide aggregate sample sizes not covered by the related adjustment table, the actual performance percentage result will be compared to the benchmark nominal percentage value. Industry aggregate performance will be identified as passing if the actual performance percentage result is greater than or equal to the benchmark nominal percentage value, and identified as failing if otherwise.

3. For CLEC-specific analysis, results with sample sizes of four or less will be aggregated into a "small sample CLEC aggregate" for each submeasure. Each small sample CLEC aggregate performance result and all remaining non-aggregated CLEC performance results will be assessed.

4. For each submeasure where the CLEC industry-wide aggregate performance *fails* the benchmark, the actual performance percentage result for each small sample CLEC aggregate and each remaining non-aggregated CLEC result will

be compared to the benchmark nominal percentage value. Each individual or aggregate performance result will be identified as passing if the actual performance percentage result is greater than or equal to the benchmark nominal percentage value, and identified as failing if otherwise.

5. For sample sizes *covered* by the related adjustment table where the CLEC industry-wide aggregate performance *passes* the benchmark, the following shall apply for each submeasure. For each benchmark submeasure, the number of performance "misses" for each small sample CLEC aggregate and each remaining non-aggregated CLEC will be compared to the number of permitted misses. CLEC performance will be identified as passing if the number of actual misses is less than or equal to the number of permitted misses, and identified as failing if otherwise.

6. For sample sizes *not covered* by the related adjustment table where the CLEC industry-wide aggregate performance *passes* the benchmark, the following shall apply. The actual performance percentage result for each small sample CLEC aggregate and each remaining non-aggregated CLEC result will be compared to the benchmark nominal percentage value. Each individual or aggregate performance result will be identified as passing if the actual performance percentage result is greater than or equal to the benchmark nominal percentage value, and identified as failing if otherwise.

Appendix D

Fisher’s Exact Test

Fisher's Exact Test

This appendix documents Fisher's Exact Test (FET) calculation methods and presents staff's comparison of Z-test and FET results.

Calculation methods

Calculation methods and examples for percentage measures where lower values represent better performance are presented in Attachment 1. Calculation methods and examples for percentage measures where higher values represent better performance are presented in Attachment 2.

Convergence of Z-test and FET results

Staff compared Type I error values (alpha probabilities) produced by the Z-test with those produced by the FET for one "lower is better" submeasure and one "higher is better" submeasure. Staff found that the results from the two tests converge for large sample sizes. Specifically, the size of the difference between the alphas calculated for each test was highly negatively correlated with the natural log of the CLEC sample size as listed in Table 1. "Highly negatively correlated" means that as sample size increases, the difference between the Z-test alpha and the FET alpha decreases in a close and predictable relationship.

Table 1

Measure type	Sample sizes	N	Correlation coefficient	p
High is better	1 to 100	102	-0.89	0.00
High is better	All	204	-0.74	0.00
Low is better	All	167	-0.94	0.00

The correlation for the whole sample for the "high is better" measure is artifactually smaller than for the half-sample because the difference between the alphas for the two tests reduced to zero and could not diminish further for very large sample sizes. Thus though the convergence was perfect for very large samples, since there was no variation, the correlation was zero for this part of the bivariate distribution.

Table 2 lists the extent of the differences between the alphas for the two tests and illustrates the convergence of the results as sample sizes increase.

Table 2

Measure type	Sample sizes	N	Mean difference	Median difference
High is better	1 to 30	63	0.12	0.09
	31 to 100	39	0.009	0.00
	101 +	102	0.0006	0.00
Low is better	1 to 100	102	0.40	0.44
	101 to 500	27	0.12	0.11
	501 to 1500	21	0.05	0.06
	1500 +	17	0.015	0.02

Mathcad worksheet: Hypothetical data example calculations for Fisher's Exact test. Measures for which low values represent good service.

Data :=

Measure	Cnum1	NC	Cval	Inum	NI	Ival
5	0	21	0.0%	6	598	1.0%
5	1	1	100.0%	231	598	38.6%
5	5	10	50.0%	321	743	43.2%
5	2	21	9.5%	234	598	39.1%
5	21	32	65.6%	321	743	43.2%
5	12	43	27.9%	345	743	46.4%
5	23	76	30.3%	321	743	43.2%
5	21	98	21.4%	210	598	35.1%
11	3	21	14.3%	32	298	10.7%
11	2	32	6.3%	98	678	14.5%
11	2	43	4.7%	76	876	8.7%
11	21	132	15.9%	98	688	14.2%
11	23	210	11.0%	101	678	14.9%
11	1	4	25.0%	8	289	2.8%
11	5	54	9.3%	6	321	1.9%
11	5	123	4.1%	32	832	3.8%
11	12	398	3.0%	34	876	3.9%
11	0	5	0.0%	0	17	0.0%
11	3	54	5.6%	7	65	10.8%
20	2	3	66.7%	65	432	15.0%
20	1	1	100.0%	210	748	28.1%
20	19	32	59.4%	154	746	20.6%
20	21	76	27.6%	111	1231	9.0%
20	3	9	33.3%	110	765	14.4%
20	5	19	26.3%	101	789	12.8%
23	0	1	0.0%	154	987	15.6%
23	1	1	100.0%	54	543	9.9%
23	3	9	33.3%	87	567	15.3%
23	2	10	20.0%	210	1122	18.7%
23	2	5	40.0%	132	876	15.1%

rows(Data) = 30
cols(Data) = 7

$HC := Data^{<1>}$ Numerator for CLEC

$NC := Data^{<2>}$ Denominator (sample size) for CLEC

$HI := Data^{<4>}$ Numerator for ILEC

$NI := Data^{<5>}$ Denominator for ILEC

The following function calculates Fisher's exact test using the above four parameters. If the CLEC numerator (HC) is zero, the probability is 1 regardless of the other parameters
.

$$FE(hc, nc, hi, ni) := \left| \begin{array}{l} x \leftarrow 1 \quad \text{if} \quad hc = 0 \\ x \leftarrow 1 - \text{phypergeom}(hc - 1, nc, ni, hi + hc) \quad \text{otherwise} \\ \text{return } x \end{array} \right.$$

$J := \text{rows}(Data) - 1$

$j := 0 .. J$

$p_j := FE\left(HC_j, NC_j, HI_j, NI_j\right)$

$Y := \text{augment}(Data, p)$

Measure	Cnum1	NC	Cval	Inum	NI	Ival	Prob
5	0	21	0.0%	6	598	1.0%	100.0%
5	1	1	100.0%	231	598	38.6%	38.7%
5	5	10	50.0%	321	743	43.2%	45.1%
5	2	21	9.5%	234	598	39.1%	100.0%
5	21	32	65.6%	321	743	43.2%	1.0%
5	12	43	27.9%	345	743	46.4%	99.5%
5	23	76	30.3%	321	743	43.2%	99.0%
5	21	98	21.4%	210	598	35.1%	99.8%
11	3	21	14.3%	32	298	10.7%	41.1%
11	2	32	6.3%	98	678	14.5%	95.5%
11	2	43	4.7%	76	876	8.7%	89.6%
11	21	132	15.9%	98	688	14.2%	35.2%
11	23	210	11.0%	101	678	14.9%	94.3%
11	1	4	25.0%	8	289	2.8%	11.8%
11	5	54	9.3%	6	321	1.9%	1.2%
11	5	123	4.1%	32	832	3.8%	53.0%
11	12	398	3.0%	34	876	3.9%	82.3%
11	0	5	0.0%	0	17	0.0%	100.0%
11	3	54	5.6%	7	65	10.8%	91.4%
20	2	3	66.7%	65	432	15.0%	6.3%
20	1	1	100.0%	210	748	28.1%	28.2%
20	19	32	59.4%	154	746	20.6%	0.0%
20	21	76	27.6%	111	1231	9.0%	0.0%
20	3	9	33.3%	110	765	14.4%	13.1%
20	5	19	26.3%	101	789	12.8%	9.1%
23	0	1	0.0%	154	987	15.6%	100.0%
23	1	1	100.0%	54	543	9.9%	10.1%
23	3	9	33.3%	87	567	15.3%	15.3%
23	2	10	20.0%	210	1122	18.7%	58.6%
23	2	5	40.0%	132	876	15.1%	16.8%

Y

Mathcad worksheet: Hypothetical data example calculations for Fisher's Exact test. Measures for which high values represent good service.

Data :=

Measure	Cnum1	NC	Cval	Inum	NI	Ival
9	1	1	100.0%	9	14	64.3%
9	3	3	100.0%	123	145	84.8%
9	6	7	85.7%	78	98	79.6%
9	9	11	81.8%	76	98	77.6%
9	14	15	93.3%	9	14	64.3%
9	17	19	89.5%	77	98	78.6%
9	17	21	81.0%	121	145	83.4%
9	23	24	95.8%	9	14	64.3%
9	24	24	100.0%	120	145	82.8%
26	145	154	94.2%	454	456	99.6%
26	276	287	96.2%	454	456	99.6%
26	321	323	99.4%	454	456	99.6%

$\mathrm{rows}(Data) = 12$
$\mathrm{cols}(Data) = 7$

$HC := Data^{\langle 2\rangle} - Data^{\langle 1\rangle}$ Numerator for CLEC. This value is converted from "hits" to "misses".

$NC := Data^{\langle 2\rangle}$ Denominator (sample size) for CLEC

$HI := Data^{\langle 5\rangle} - Data^{\langle 4\rangle}$ Numerator for ILEC, also converted from "hits" to "misses."

$NI := Data^{\langle 5\rangle}$ Denominator for ILEC

The following function calculates Fisher's exact test using the above four parameters. If the CLEC numerator (HC) is zero, the probability is 1 regardless of the other parameters.

$$FE(hc,nc,hi,ni) := \left| \begin{array}{l} x \leftarrow 1 \quad \text{if} \quad hc = 0 \\ x \leftarrow 1 - \mathrm{phypergeom}(hc-1, nc, ni, hi+hc) \quad \text{otherwise} \\ \text{return } x \end{array} \right.$$

$J := \mathrm{rows}(Data) - 1$

$j := 0..J$

$p_j := FE(HC_j, NC_j, HI_j, NI_j)$

$Y := augment(Data, p)$

Measure	Cnum1	NC	Cval	Inum	NI	Ival	Prob
9	1	1	100.0%	9	14	64.3%	100.0%
9	3	3	100.0%	123	145	84.8%	100.0%
9	6	7	85.7%	78	98	79.6%	80.1%
9	9	11	81.8%	76	98	77.6%	74.9%
9	14	15	93.3%	9	14	64.3%	99.4%
9	17	19	89.5%	77	98	78.6%	93.0%
9	17	21	81.0%	121	145	83.4%	48.9%
9	23	24	95.8%	9	14	64.3%	99.9%
9	24	24	100.0%	120	145	82.8%	100.0%
26	145	154	94.2%	454	456	99.6%	0.0%
26	276	287	96.2%	454	456	99.6%	0.1%
26	321	323	99.4%	454	456	99.6%	55.0%

Y

Appendix E

Binomial Exact Test

Binomial Exact Test

This appendix documents binomial exact test calculation methods and presents staff's comparison of Z-test and binomial test results. Calculation methods and examples for rate measures are presented in Attachment 1.

Convergence of Z-test and binomial exact test results

Staff compared Type I error values (alpha probabilities) produced by the Z-test with those produced by the binomial test for submeasure. As with the Fisher's Exact Test, staff found that the results from the two tests converge for large sample sizes. Specifically, the size of the difference between the alphas calculated for each test was highly negatively correlated with the natural log of the CLEC sample size as listed in Table 1. "Highly negatively correlated" means that as sample size increases, the difference between the Z-test alpha and the binomial test alpha decreases in a close and predictable relationship.

Table 1

N	Correlation coefficient	p
117	-0.93	0.00

Table 2 lists the extent of the differences between the alphas for the two tests and illustrates the convergence of the results for the two tests.

Table 2

Sample sizes	N	Mean difference	Median difference
1 to 100	61	0.32	0.38
101 to 300	37	0.05	0.05
300 +	19	0.008	0.00

Excell spreadsheet formula for binomial exact test calculations

The Excell© worksheet cell entry that calculates alpha for the binomial exact test is as follows:

=1-IF(B2=0,0,BINOMDIST(B2-1, B2+E2, C2/(C2+F2),TRUE))

Using trouble report rates as an example rate performance measure, column B contains “Cnum1,” the number of CLEC troubles; column C contains “N_c” the number of lines, column E contains “Inum,” the number of ILEC troubles; and column F contains N_i, the number of ILEC lines. The above formula is the cell entry for the first row of performance results in the spreadsheet (row 2) presented on the next page. The data is hypothetical data for demonstration purposes only.

Excell spreadsheet: Hypothetical data example of binomial exact test calculations.

	A	B	C	D	E	F	G	H
1	Measure	Cnum1	N_c	Cval	Inum	N_i	Ival	α
2	15	0	143	0.00%	987	1876543	0.05%	1.00
3	15	3	343	0.86%	4321	2012345	0.20%	0.04
4	15	1	432	0.22%	1321	2012345	0.07%	0.25
5	15	4	876	0.45%	4321	2012345	0.20%	0.12
6	15	2	2987	0.07%	3210	2101234	0.15%	0.94
7	15	6	4321	0.14%	2432	2101234	0.11%	0.38
8	15	5	5432	0.08%	2765	1876543	0.15%	0.90
9	15	7	13210	0.05%	1765	2012345	0.09%	0.94
10	15	8	13210	0.06%	4321	2012345	0.20%	1.00
11	16	0	4	0.00%	32	14321	0.21%	1.00
12	16	3	12	25.00%	876	7654	10.66%	0.16
13	16	2	13	15.38%	987	43210	2.20%	0.04
14	16	8	21	40.00%	876	7654	10.66%	0.00
15	16	1	21	4.55%	1231	48765	2.56%	0.41
16	16	3	76	3.90%	876	7654	10.66%	0.99
17	16	9	98	9.38%	543	21012	2.65%	0.00
18	16	6	132	4.62%	12101	543210	2.32%	0.08
19	16	7	187	3.83%	8987	432101	1.96%	0.10
20	16	4	198	2.06%	10123	498765	2.12%	0.57
21	16	5	365	1.39%	11012	454321	2.45%	0.94
22	19	0	1	0.00%	2799	54321	4.91%	1.00
23	19	2	18	11.11%	1012	321012	0.35%	0.00
24	19	1	54	1.82%	1012	321012	0.34%	0.16
25	19	8	54	13.56%	2987	65432	4.89%	0.00
26	19	7	87	7.95%	2987	65432	4.86%	0.11
27	19	0	87	0.00%	26543	3432101	0.72%	1.00
28	19	5	321	1.61%	987	301234	0.31%	0.00
29	19	9	876	1.09%	1876	210123	0.90%	0.38
30	19	4	987	0.44%	26543	3654321	0.72%	0.93
31	19	6	1210	0.47%	143210	12345678	1.34%	0.99

Appendix F

Beta Error

Beta Error Levels

This appendix documents staff's analyses of beta error levels for various performance. Staff prepared two analyses. The first analysis examined betas for all possible parity measures using Modified Z-test calculations for all measure types. While these are not the test applications that the Commission will implement, using these tests allows some comparisons that are otherwise difficult. These values are calculated from May 2000 performance data. The alternative hypothesis posed for all estimates was that the CLEC's results were at least 50 percent worse than the ILEC's results. The formula used is based on Hays, *supra* at 284-289 (1994) except that the ILEC and CLEC sample sizes are used:

$$t_{\beta} = (H_0 - H_{alt}) / SD_m$$

The second analysis examined beta error levels for all parity measures as implemented by the Commission in this decision with the exception that log transformations were not performed. These values are calculated from July through September, 2000 performance data. The above formula was used for the average-based parity measures. Pacific's Dr. Gleason calculated the betas for the percentage and rate measures using the hypergeometric and binomial distributions, respectively.

Table 1 lists beta values calculated from Pacific's May, 2000 performance data as described above. Calculations are presented for four different critical alpha levels and for two alternative hypotheses. The alternative hypotheses represents performance provided to CLECs that is 50 percent worse (150%) and 100 percent worse (200%) than performance the ILEC provides itself. For example, the mean beta value for a critical alpha level of 0.10, given an alternative hypothesis of 50 percent worse performance, is 0.63. This should be interpreted as: If we keep Type I error to a maximum of 10 percent (α_{crit} = 0.10), on average we will experience a 63-percent error rate when trying to detect performance for the CLEC that is at least 50 percent worse than performance for the ILEC.

Table 1

Average Beta values for Pacific May, 2000, parity measures				
	Alternative hypothesis			
	150%		200%	
Critical α	Mean	Median	Mean	Median
0.05	0.70	0.88	0.58	0.77
0.10	0.63	0.79	0.51	0.64
0.15	0.57	0.72	0.45	0.55
0.20	0.52	0.65	0.40	0.47

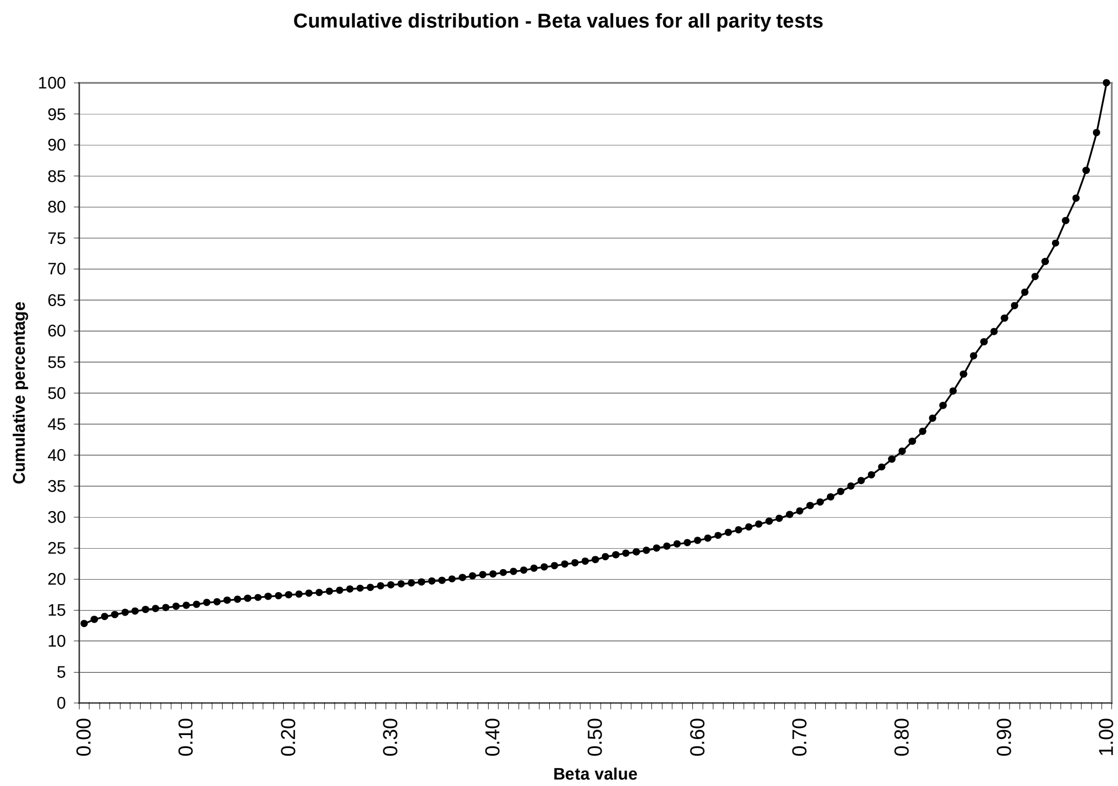
Table 2 presents beta values for parity measures by measurement type for Pacific's performance in July through September, 2000. All beta calculations are based on a 0.10 critical alpha and an alternative hypothesis of 50 percent worse performance (150%) for CLECs.

Table 2

Average beta value by parity test type - Pacific performance July through August, 2000					
	All	Average	Percent (Hi)	Percent (Lo)	Rate
N	9909	2768	928	3558	2655
Percentage	100%	28%	9%	36%	27%
Mean	**0.70**	**0.45**	**0.42**	**0.87**	**0.83**
Median	**0.85**	**0.58**	**0.36**	**0.94**	**0.92**
SD	0.35	0.37	0.41	0.20	0.24
Skewness	-1.15	-0.20	0.18	-2.52	-2.18
Kurtosis	-0.24	-1.76	-1.75	6.58	4.16
Minimum	0.00	0.00	0.00	0.00	0.00
Maximum	1.00	0.89	1.00	1.00	1.00

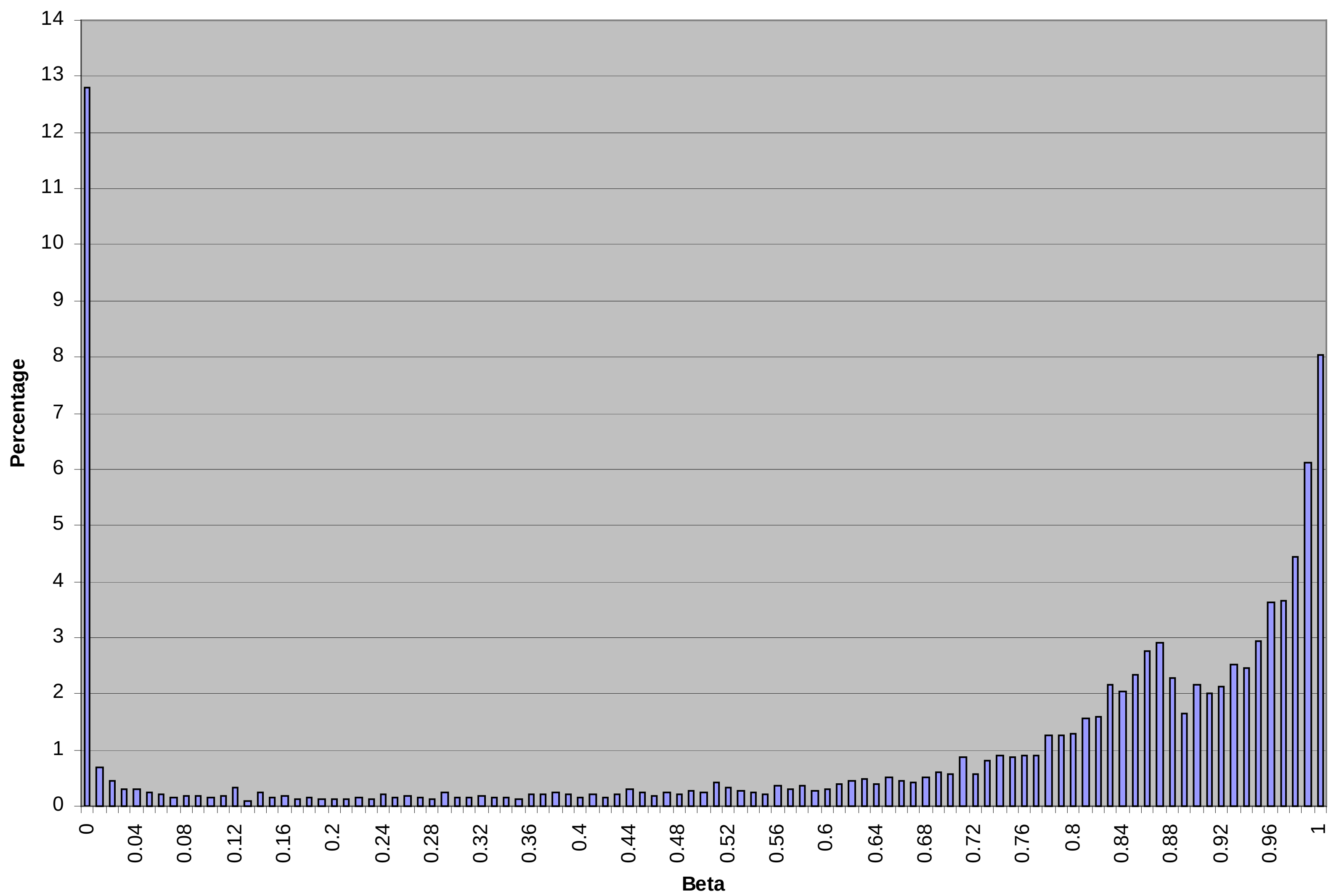
Attachment 1 presents the beta value cumulative distribution for all parity measures as presented in Table 2. For example, about 16 percent of all CLEC submeasure parity test results have beta values of 0.10 or less. In other words, when Type I error rate is held to 0.10 or less for all results, only 16 percent of all parity test results will have a Type II error rate of 0.10 or less.

Attachment 2 presents the beta value frequency distributions for all parity measures combined and each measure type as presented in Table 2. For example, for all parity measures, two percent of beta values equal 0.84 (page 1).

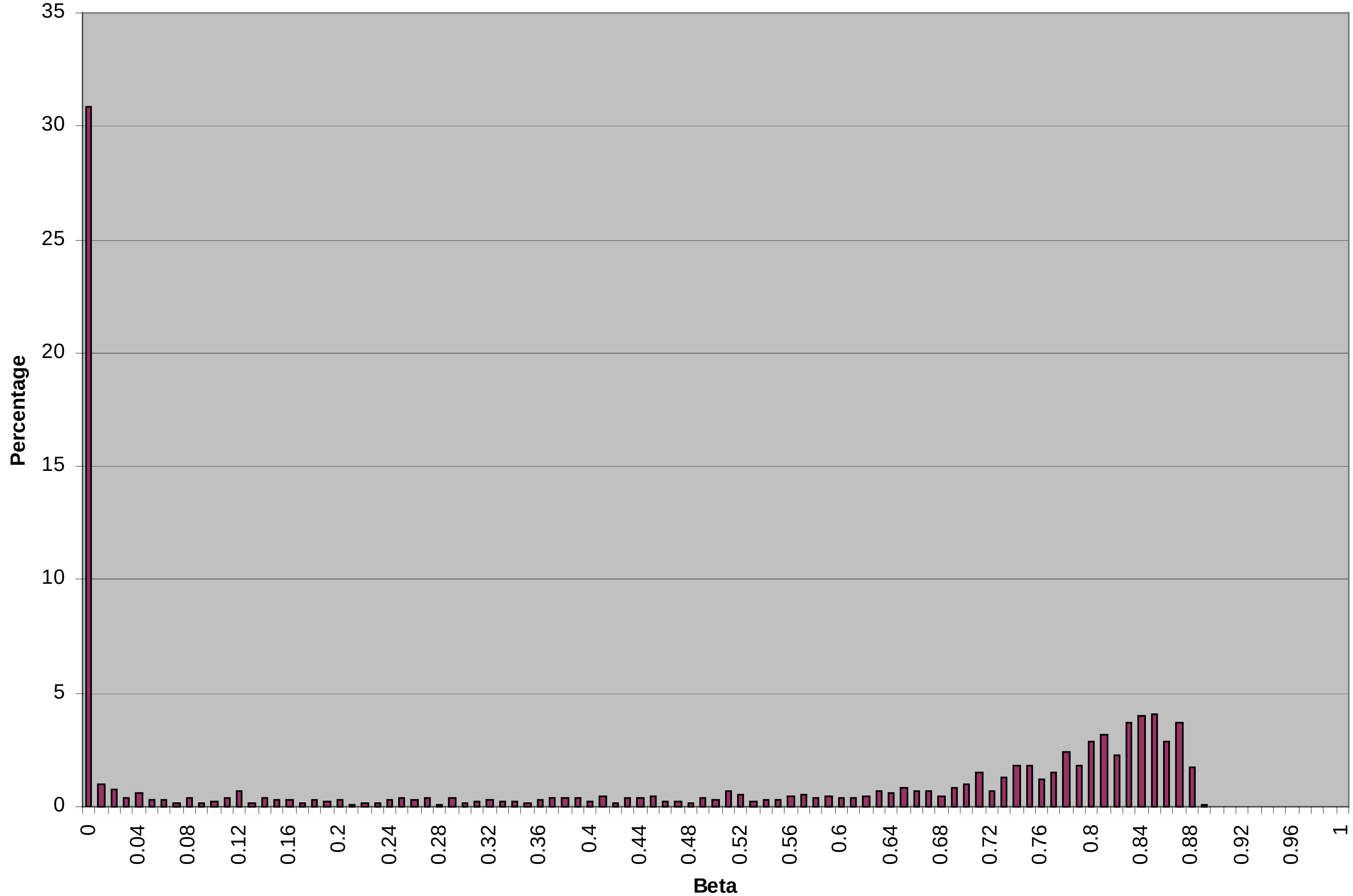


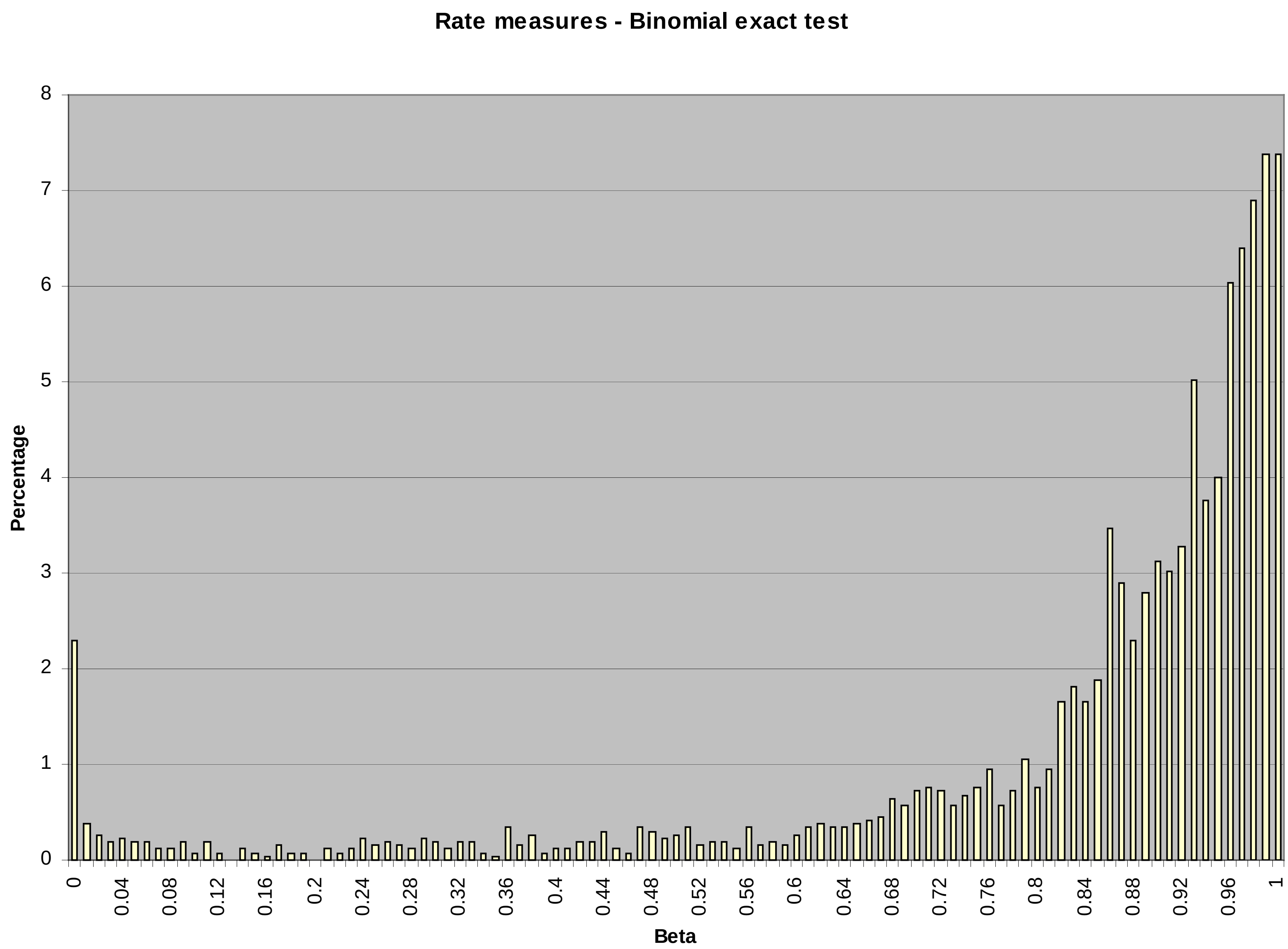
Cumulative distribution - Beta values for all parity tests
Cumulative percentage
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
Beta value

All measures

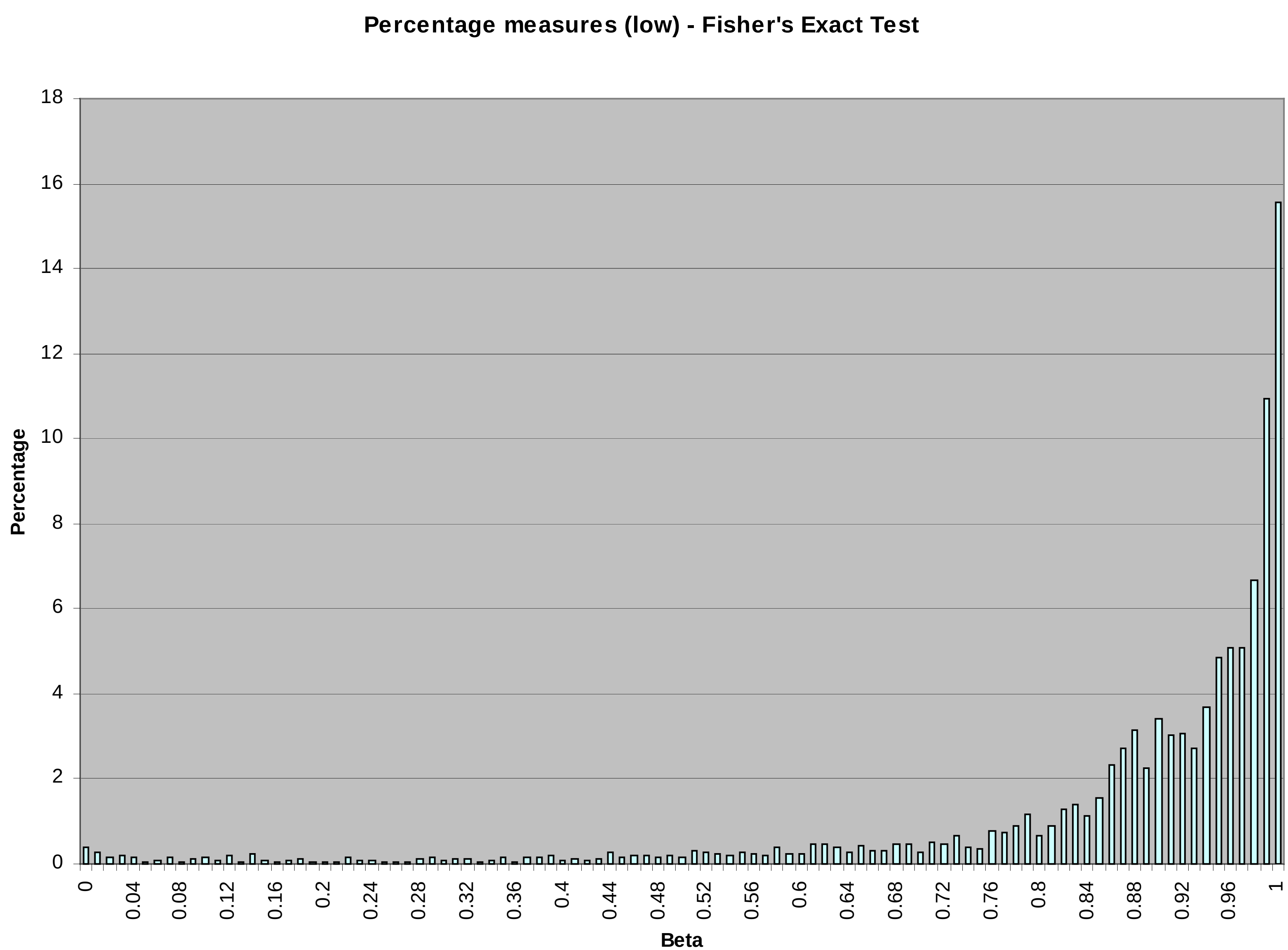


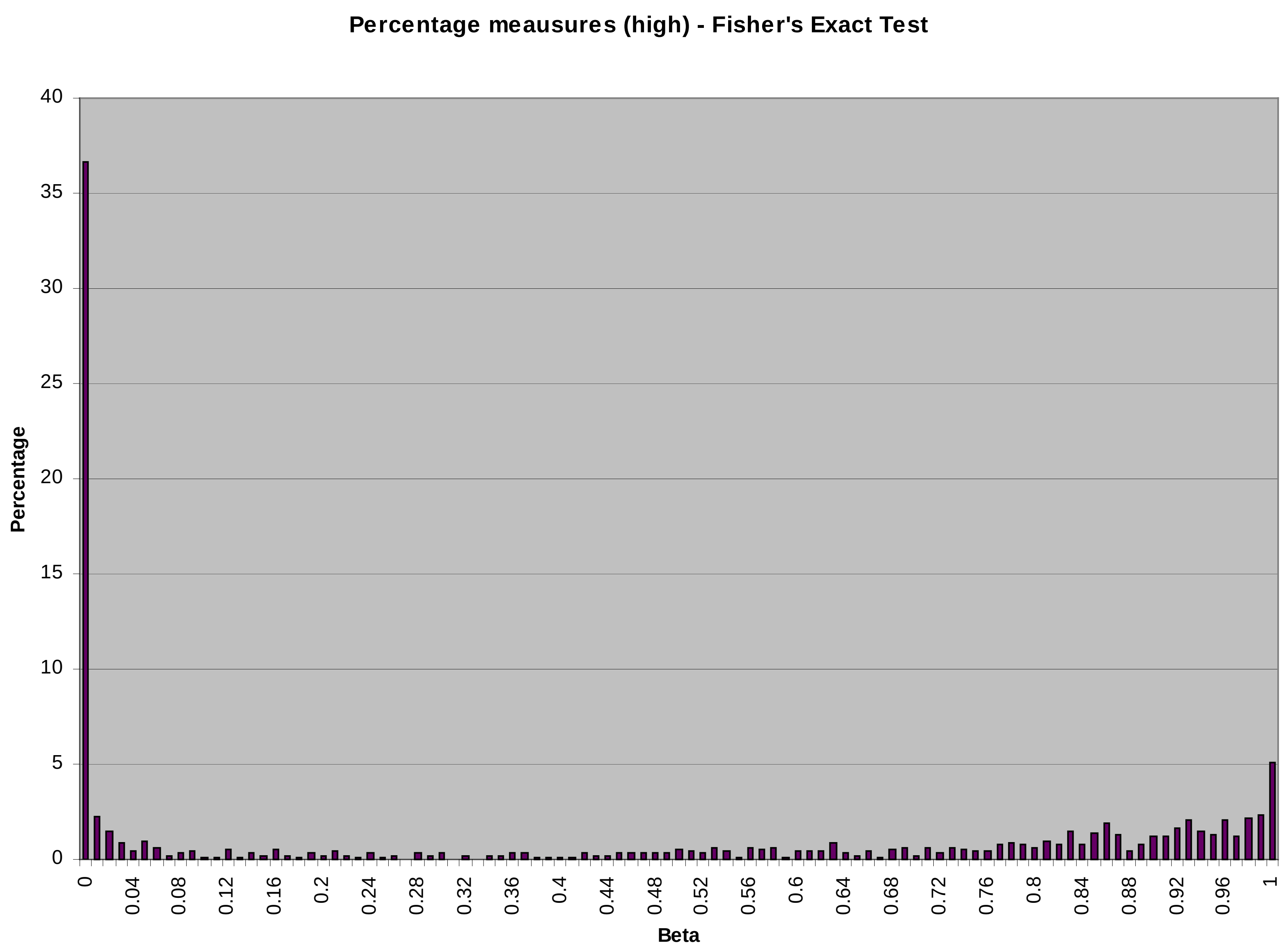
Average measures - Modified t-test (no transformation)





Rate measures - Binomial exact test
Percentage
0
1
2
3
4
5
6
7
8
Beta
0
0.04
0.08
0.12
0.16
0.2
0.24
0.28
0.32
0.36
0.4
0.44
0.48
0.52
0.56
0.6
0.64
0.68
0.72
0.76
0.8
0.84
0.88
0.92
0.96
1


Percentage measures (low) - Fisher's Exact Test
Percentage
0
2
4
6
8
10
12
14
16
18
Beta
0
0.04
0.08
0.12
0.16
0.2
0.24
0.28
0.32
0.36
0.4
0.44
0.48
0.52
0.56
0.6
0.64
0.68
0.72
0.76
0.8
0.84
0.88
0.92
0.96
1


Percentage meausures (high) - Fisher's Exact Test
Percentage
Beta
0
5
10
15
20
25
30
35
40
0
0.04
0.08
0.12
0.16
0.2
0.24
0.28
0.32
0.36
0.4
0.44
0.48
0.52
0.56
0.6
0.64
0.68
0.72
0.76
0.8
0.84
0.88
0.92
0.96
1

Appendix G

Balancing Alpha and Beta Error

Balancing Alpha and Beta Error

This appendix documents staff's efforts to balance alpha and beta error levels for performance result assessment. To calculate the "balance point" for alpha and beta error staff adapted a balancing formula presented in Das (1994). Staff treated this formula as a "equal error" formula by assuming equal consequences for the two types of error. The formula was also adapted by including the "N" for both ILEC and CLEC samples. The formula as used was:

$Z_{\beta} = ((H_0 - H_{alt}) / SD_i * \text{sqrt}(1 / N_c + 1 / N_i)) / 2$

Where:
H_0 = Null hypothesis (ILEC mean)
H_{alt} = Alternate hypothesis
SD_I = ILEC standard deviation
N_c = CLEC sample size
N_I = ILEC sample size

Staff analyzed Pacific's May, 2000, performance results to estimate the effects of setting critical alpha levels equal to beta error for each result. An alternate hypothesis of 50-percent worse performance was assumed for the calculations. In other words, staff estimated the critical alpha level that would result in equal error (beta) in detecting performance at least 50% worse for the CLEC as for the ILEC. On the average, alpha balanced with beta at a value of 0.33. In other words, if alpha error was held to a maximum of 33 percent, beta error would also be 33 percent. Table 1 presents the summary statistics.

Table 1

	Alpha balanced with beta
N	3481
Mean	0.33
Median	0.41
Minimum	0
Maximum	0.5

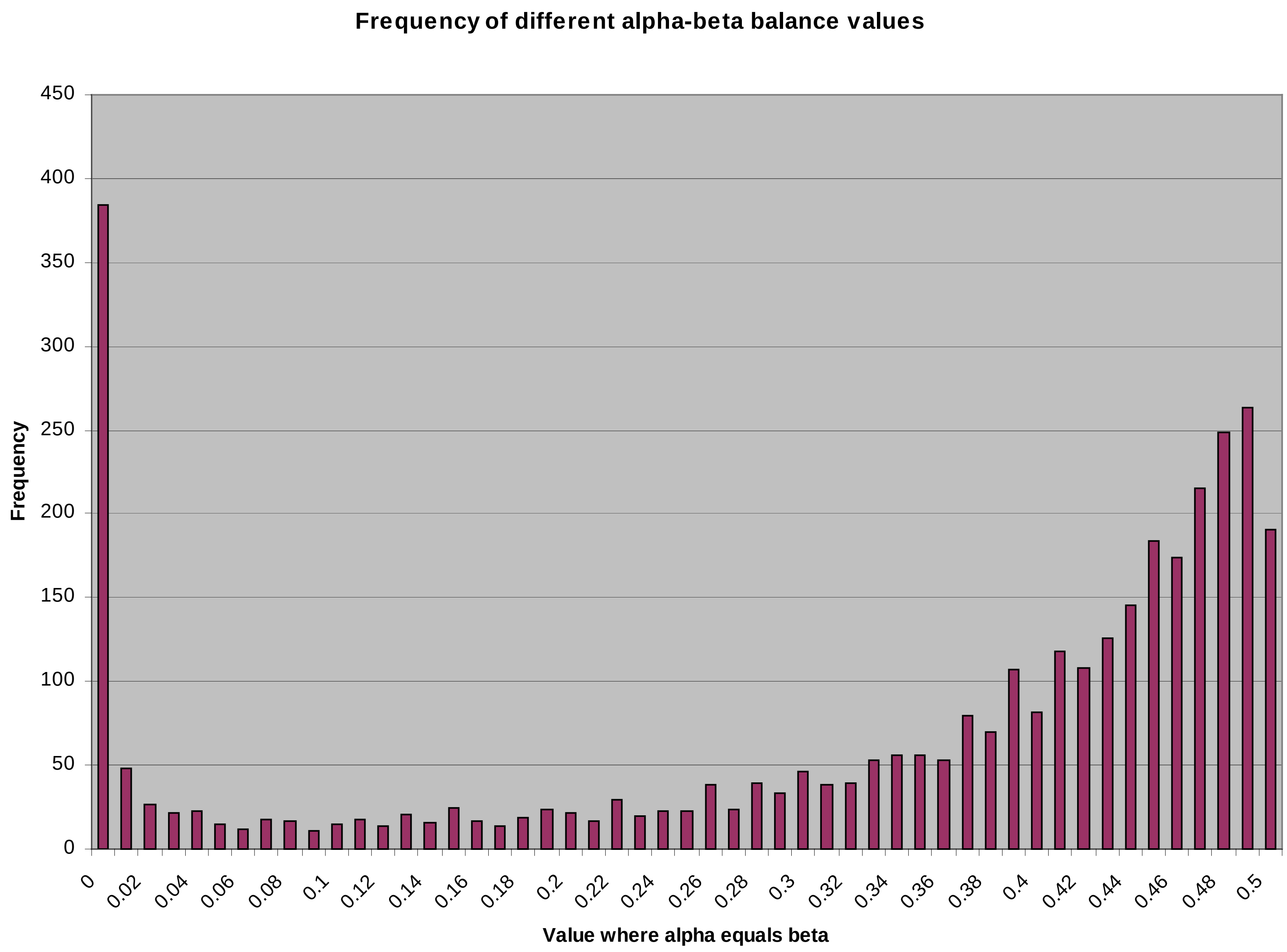
Attachment 1 presents a frequency distribution of the balancing values.

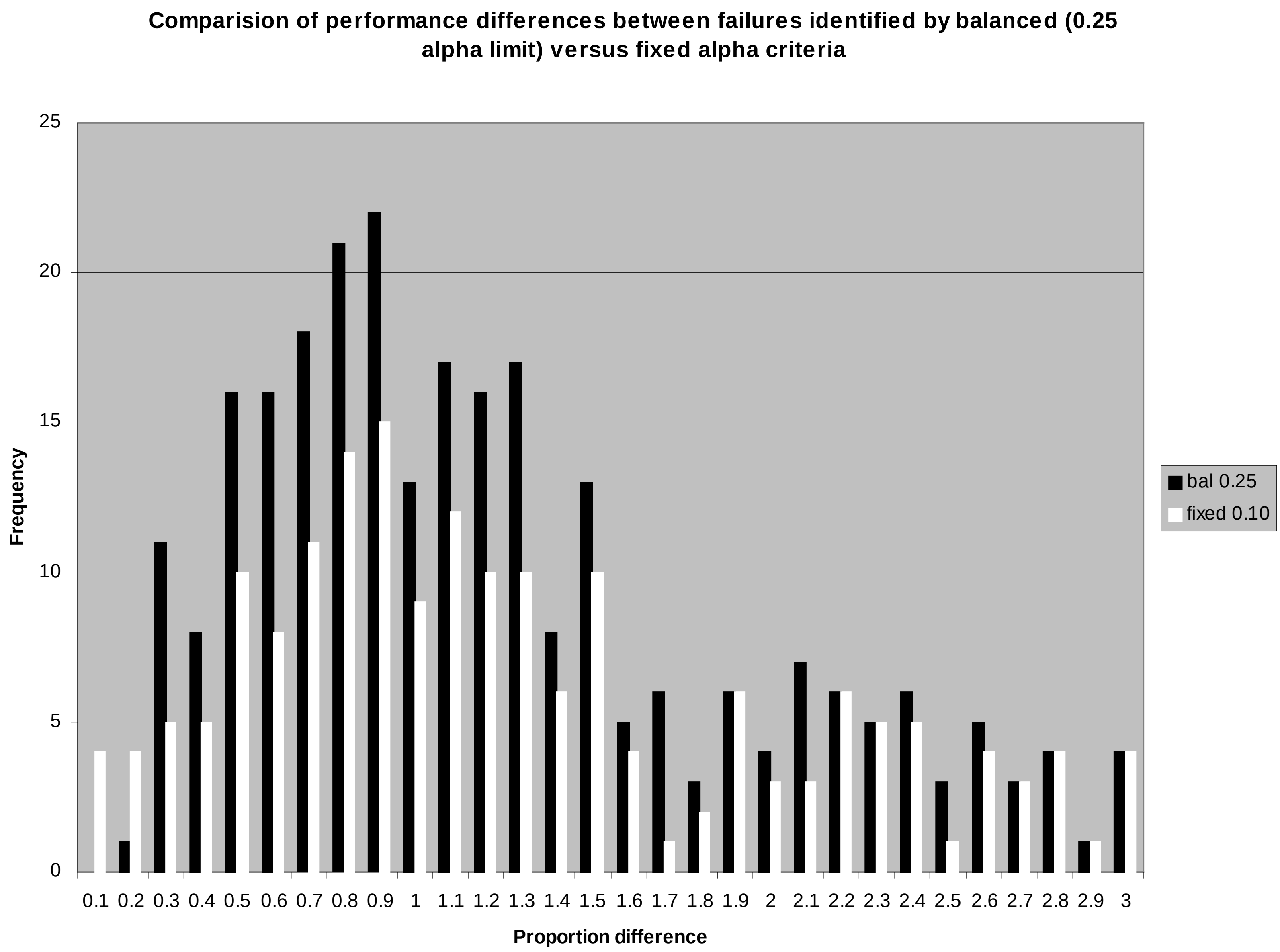
Staff also calculated the resulting error rates with an alpha error rate "ceiling." Table 2 presents summary statistics for those calculations.

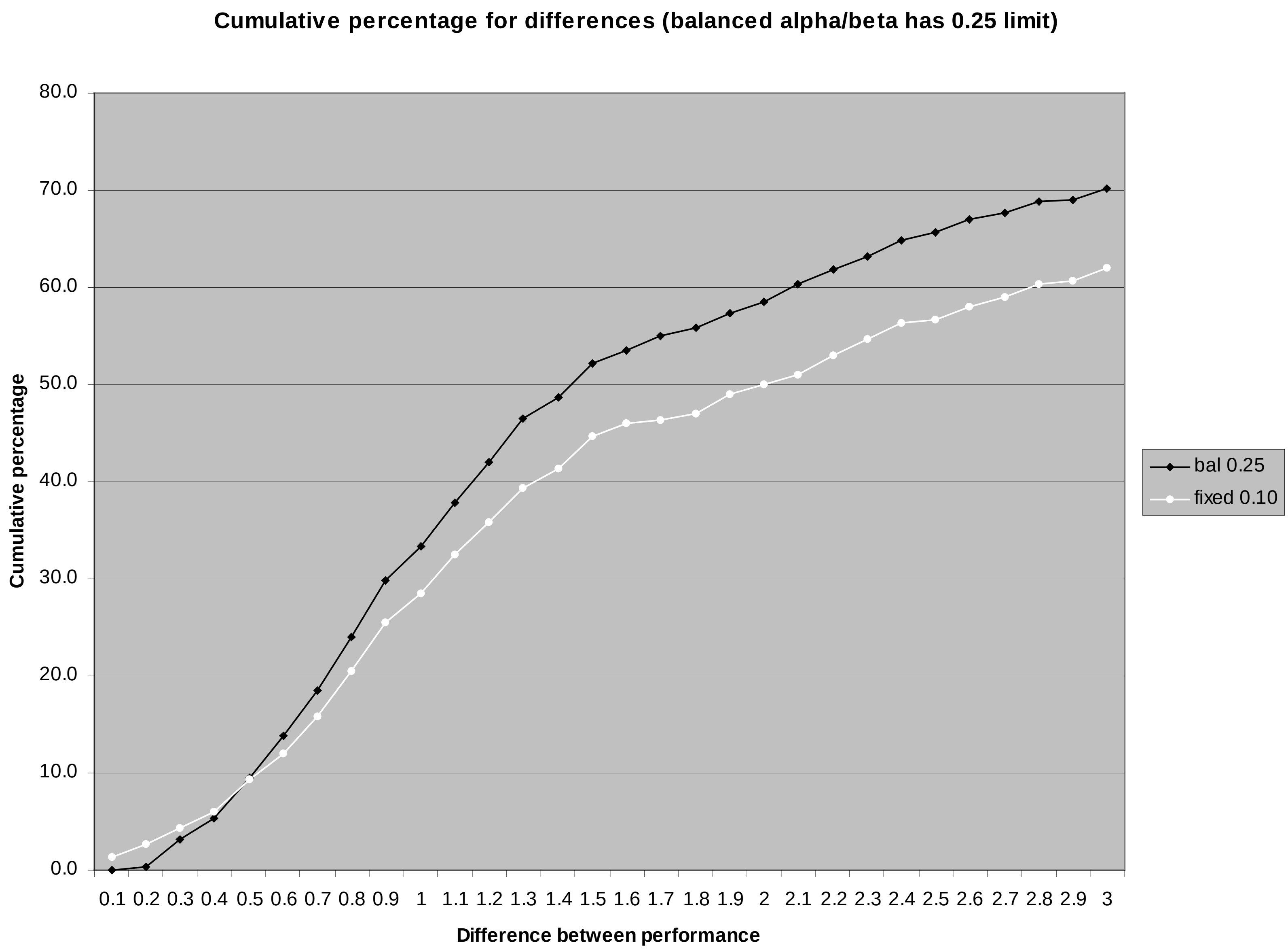
Table 2

Critical alpha levels resulting from different alpha/beta balance limits			
	Alpha limit		
	0.33	0.25	0.2
N	1204	894	782
Mean	**0.131**	**0.0726**	**0.0499**
Median	**0.11**	**0.02**	**0.01**
Mode	0	0	0
Minimum	0	0	0
Maximum	0.33	0.25	0.2

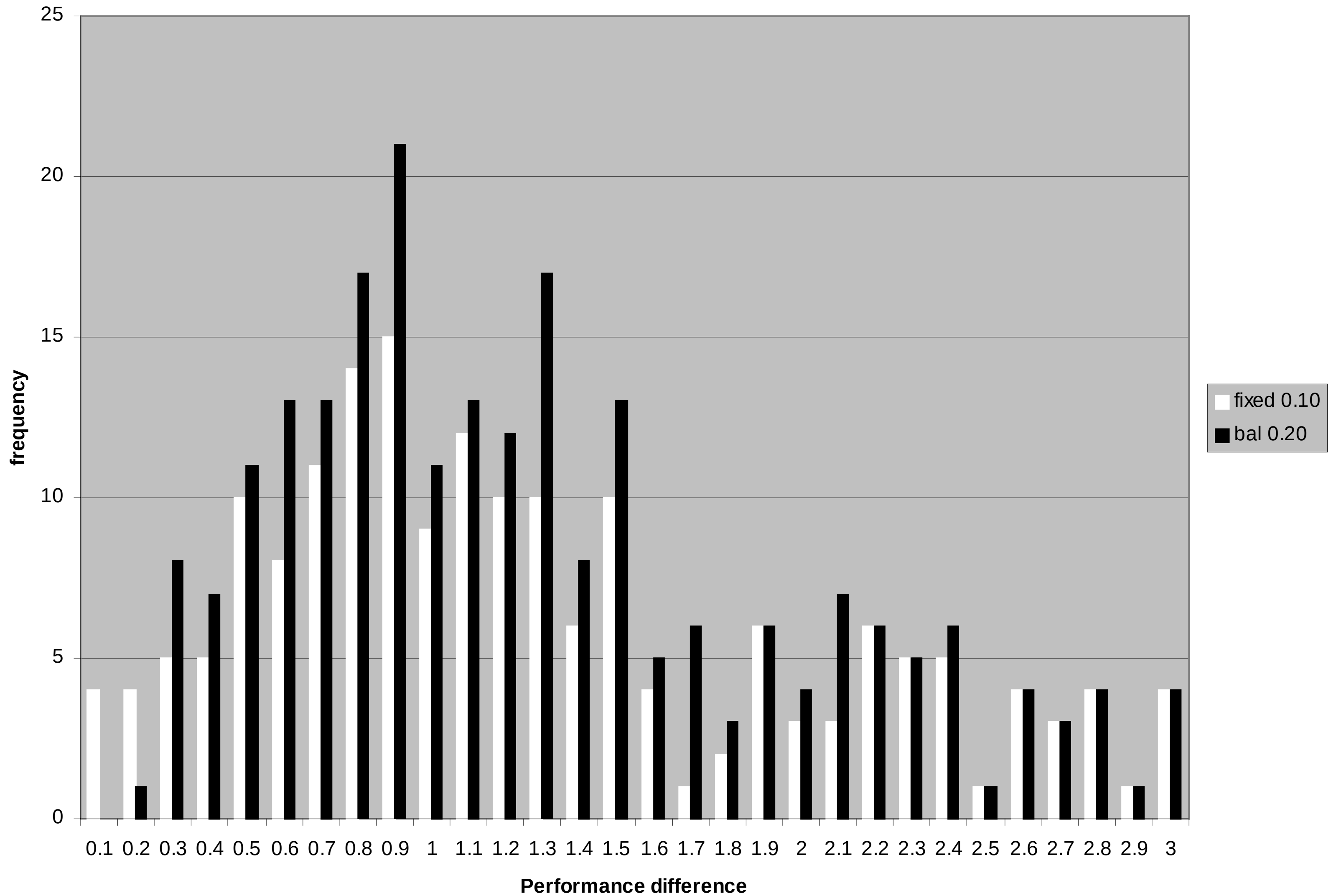
Staff also examined the net effect on the size of the difference between ILEC and CLEC performance that would be identified as a performance failure. Theoretically, balancing alpha and beta should result in an increase in larger differences being detected and a decrease in smaller differences being detected. Attachment 2 shows that this in fact would occur. Limiting alpha to 0.25, for example, results in a lower proportion of failure identifications where performance to a CLEC is zero to 50 percent worse than ILEC performance to itself (Attachment 2, page 2), relative to a fixed 0.10 alpha criterion. Conversely, this limit results in a greater proportion of failure identifications where performance to a CLEC is *at least* 50 percent worse than ILEC performance to itself. These charts only display results up to the point where performance to a CLEC is three times worse than performance for the ILEC. At this point, however, there are no further differences between a fixed 0.10 alpha criterion and either the 0.20 or 0.25 alpha / beta balance limited criteria.

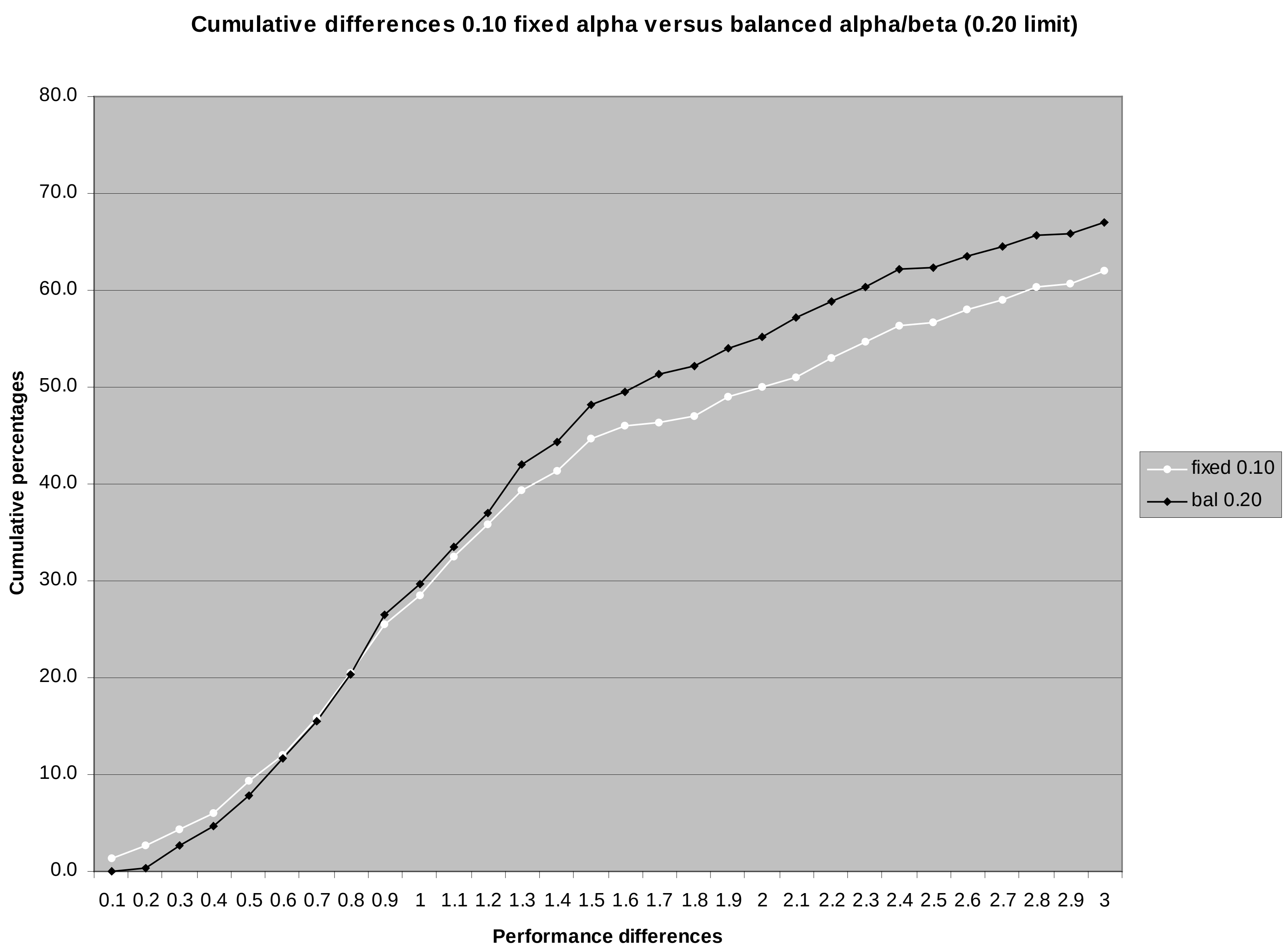

Frequency of different alpha-beta balance values
Frequency
0
50
100
150
200
250
300
350
400
450
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
0.22
0.24
0.26
0.28
0.3
0.32
0.34
0.36
0.38
0.4
0.42
0.44
0.46
0.48
0.5
Value where alpha equals beta


Comparision of performance differences between failures identified by balanced (0.25 alpha limit) versus fixed alpha criteria
Frequency
0
5
10
15
20
25
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3
Proportion difference
bal 0.25
fixed 0.10


Cumulative percentage for differences (balanced alpha/beta has 0.25 limit)
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Cumulative percentage
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3
Difference between performance
bal 0.25
fixed 0.10

Frequency differences between fixed 0.10 alpha and balanced alpha/beta (0.20 limit)





Cumulative differences 0.10 fixed alpha versus balanced alpha/beta (0.20 limit)
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
Cumulative percentages
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3
Performance differences
fixed 0.10
bal 0.20

Appendix H

Pacific’s Proposed Aggregation Rules

Effects of Pacific's proposed aggregation rules

Month	Total number of parity results	Number of results with samples larger than 10		Number of results with samples smaller than 10	Number of small samples that aggregate to samples 5 or larger		Number of small samples that aggregate to samples smaller than 5	Number of industry aggregate samples 5 or larger		Number of industry aggregate samples smaller than 5 (discarded)	
January	3062	1719	56.1%	1343	1221	39.9%	122	36	1.2%	86	2.8%
February	3138	1795	57.2%	1343	1257	40.1%	86	30	1.0%	56	1.8%
March	3425	1950	56.9%	1475	1348	39.4%	127	45	1.3%	82	2.4%
Total	9625	5464	56.8%	4161	3826	39.8%	335	111	1.2%	224	2.3%

Note: This table presents aggregation rules results for parity measures only.

Appendix H: Pacific's proposed aggregation rules

Appendix H, Attachment 1

SAMPLE SIZE AGREEMENT

Submitted to ALJ Reed on 4/25/00

RULES: Only applicable to sub-measures that would normally have small sample sizes for all CLECs, ie., the process etc., being measured isn't something that is generally ordered a lot in a month.

The following measures and sub-measures are not subject to minimum sample size. Data for the following will not be discarded, but rather incentives will apply, once incentives are ordered.

What is agreed to in this memo is subject to appropriate incentives review, when ordered.

Measure 30: Agreed to by Pacific Bell and GTEC

Measure 40: Agreed to by Pacific Bell and GTEC

Measure 41: Agreed to by Pacific Bell and GTEC

UNE Loop DS-3: (Disaggregated as an Service Group Type) Agreed to by Pacific Bell, GTEC checking, but no GTEC commitment yet.

UNE-Transport DS-1: (Disaggregated within UNE-Transport) Agreed to by Pacific Bell, GTEC checking, but no GTEC commitment yet.

UNE-Transport DS-3: (Disaggregated within UNE-Transport) Agreed to by Pacific Bell, GTEC checking, but no GTEC commitment yet.

Interconnection Trunks: Agreed to by Pacific Bell and GTEC

Note: OC level services will also be added to this list if agreed to as a service group type as part of the JPSA performance measurements.

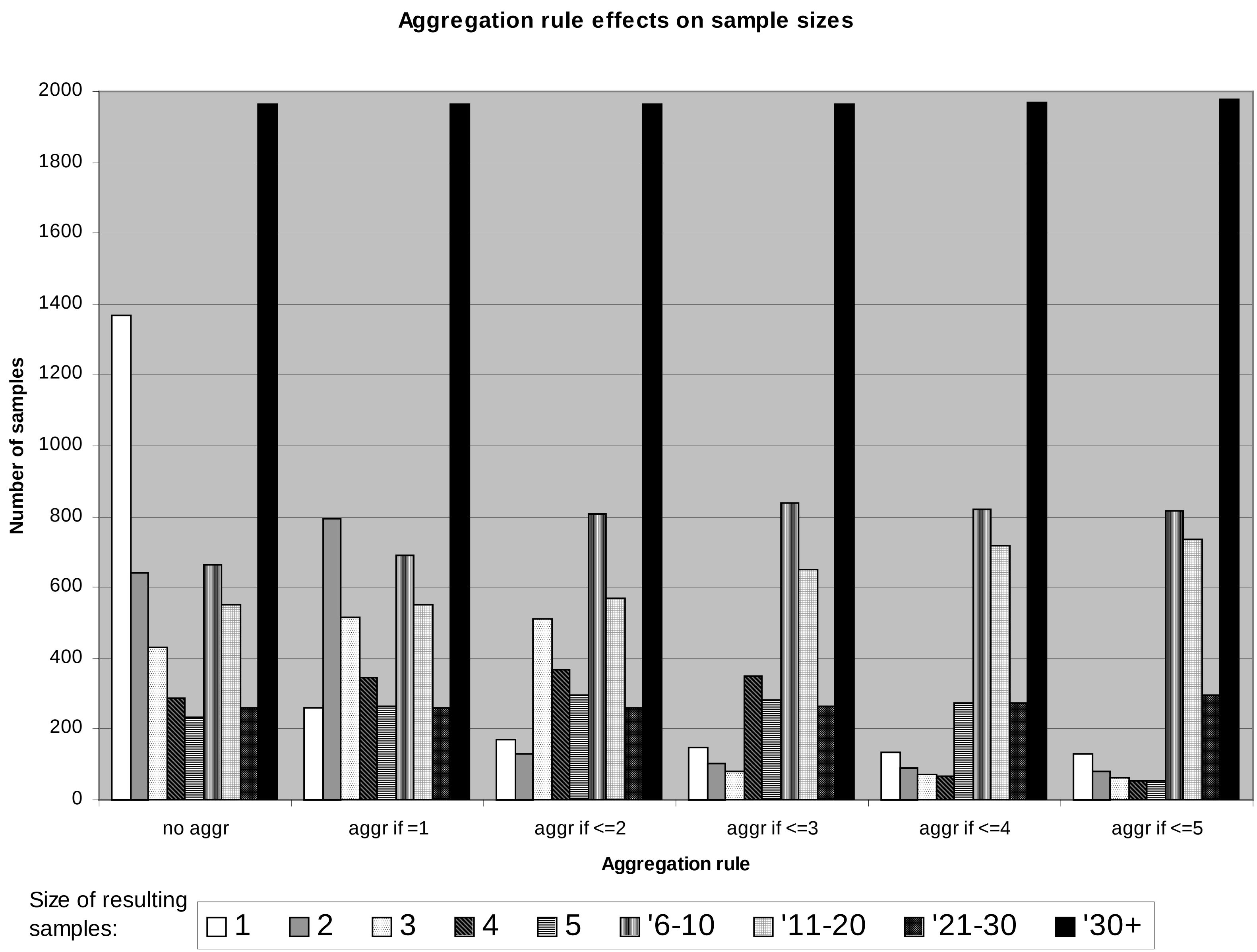
Appendix I

Implemented Aggregation Rule Results

Aggregation Rule Proposal Results

This appendix illustrates the effects of the aggregation rules implemented by this Decision. The attached chart shows how the frequency of small samples is diminished as increasingly larger sample sizes are subject to aggregation. For example, all possible aggregations of sample sizes of one reduces the number of sample sizes of one from nearly 1400 samples (see the white bar for the bar cluster labeled "no aggr") to approximately 250 samples (see the white bar for the bar cluster labeled "aggr if = 1"). The bar cluster labeled "aggr if <= 4" represent the final results of the implemented aggregation rules.

These aggregation rules avoid some of the potential pitfalls of the previously proposed rules as discussed in the body of the Decision. First, small sample results are only aggregated with like small sample results, thus minimizing the likelihood that larger sample results will "mask" the small sample results. Second, while small sample failures may still cause non-failing samples to fail when aggregated, the incentive phase of the proceeding can address the problem of potential spurious allocation of incentive payments. Third, unnecessary aggregation is minimized. No small samples are aggregated with large samples as only like-size samples are aggregated. Fourth, since only like-sized small samples are aggregated, there is no ambiguity about which results, aggregated or non-aggregated, should determine results.


Aggregation rule effects on sample sizes
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Number of samples
no aggr
aggr if =1
aggr if <=2
aggr if <=3
aggr if <=4
aggr if <=5
Aggregation rule
Size of resulting samples:
1
2
3
4
5
'6-10
'11-20
'21-30
'30+

Appendix J

Log transformations

Introduction

This appendix documents staff's inquiry into the problem of applying distribution-based statistical tests to average-based performance measures. Unlike for percentage-based and rate-based measures, no distribution-free tests for average-based measures are ready to implement given the current record in this proceeding. Consequently, the only current test option is the modified *t*-test. Staff's primary concern is that the accuracy of normal distribution based statistical tests, such as the *t*-test, diminishes for smaller samples to the degree that those samples depart from normality.[1] The Central Limit Theorem states that with larger samples, the sampling distribution of means is normally distributed even for non-normal raw score distributions. However, the degree of non-normality, and especially the degree of asymmetry, affects the sampling distribution, and especially affects one-tailed tests.[2]

Staff investigated data transformations for applying normal distribution based tests to non-normal data.[3] The investigation examines: (1) several actual performance data distributions, (2) a theoretical sampling mean distribution, (3) the statistical effects of several data-normalizing transformations, and (4) the performance evaluation implications of the most statistically appropriate transformation. Conclusions regarding the best option are presented.

Method and Results

Performance result distributions

Sixteen average-based ILEC performance submeasure distributions from one performance measure were examined.[4] Statistics for non-normality, skewness and kurtosis, were calculated. All but one of the distributions were positively skewed[5] and all but two were leptokurtic.[6] While a normal

---

[1] Winer (1971), p. 6.

2 Hays (1997), pp. 327-328; McNemar (1962), pp. 106-107.

[3] Performance measure 34 does not measure time to complete a task and is likely normally distributed. It is excluded from this discussion and will not be subject to log transformation. *See* Pacific Bell Comments on the Draft Decision at 3 (December 18, 2000).

[4] These distributions were provided by Pacific Bell. Staff also examined six distributions provided by Verizon and found the shape of those distributions to be similar. Those results are not reported here.

[5] Positive values indicate positive skewness, that is, the observations are concentrated at the lower end of the scale and gradually trail off to fewer observations in a longer "tail" to the right, the higher part of the

Footnote continued on next page

curve has skeweness and kurtosis values of zero, the skewness of the sixteen submeasures ranged from 0.0 to 28.3, and the kurtosis ranged from –1.6 to 1746. Table 1 reports the skewness and kurtosis for the sixteen distributions. The frequency distribution graphs for these sixteen distributions are presented in Attachment 1 to this appendix.

Table 1

Submeasure	N	Mean	Median	Skewness	Kurtosis
Ex. 1	179254	1.18	0	21.0	1430.3
Ex. 2	23608	1.60	0	15.4	503.0
Ex. 3	19943	6.91	6	12.1	271.2
Ex. 4	17951	0.92	0	28.3	1745.7
Ex. 5	17940	2.76	2	15.6	590.3
Ex. 6	11864	1.40	0	9.1	184.6
Ex. 7	9149	1.29	0	19.7	661.2
Ex. 8	6827	2.48	1	10.3	198.6
Ex. 9	6340	3.05	1	5.3	48.2
Ex. 10	771	8.18	7	6.9	105.3
Ex. 11	538	7.89	7	1.8	8.2
Ex. 12	34	71.62	20	0.5	-1.6
Ex. 13	14	34.36	20.5	1.4	1.5
Ex. 14	9	6.00	4	1.9	4.0
Ex. 15	8	47.50	40.5	0.7	-0.1
Ex. 16	6	10.50	10.5	0.0	2.1

Academic theory indicates data from measures of time to complete a task are lognormally distributed.[7] Overall, the skewness and kurtosis of these distributions were consistent with what would be expected from a lognormal distribution. Only three of the five smallest samples (n < 35) had skewness less than one (< 1). Only two of the smallest five samples had kurtosis less than one (< 1).

---

scale. Briefly stated, a positively skewed distribution has a longer tail for higher scores than for lower scores. Negative skewness values indicate the reverse, that is, a longer tail for the lower scores relative to the higher scores.

[6] Positive values indicate leptokurtic distributions, that is, the distribution is more peaked than a normal distribution. Negative values indicate platykurtic distributions, that is, the distribution is flatter than a normal distribution.

[7] Winer (1971), p. 400.

Theoretical sampling mean distributions

To examine the extent of the problem posed by the skewness of the data, simulated distributions were examined to investigate the sample sizes necessary to achieve normality in the sampling mean distribution. While Central Limit Theorem poses that sampling mean distributions will be normal for many large non-normal samples, it is not clear if Central Limit Theorem's general tenet applies to the data for these measures.

At staff's request, Pacific Bell's consultant, Dr. Gleason, created a MathCad© worksheet to generate multiple samples from a lognormal distribution. This worksheet is included as Attachment 2. Using this worksheet five analyses were repeated for several selected performance results. The analysis summary following the worksheet pages shows that even for samples as large as 1000, many distributions are non-normal, the degree of departure from normality can be highly variable, and the log transformation notably improves normality.

Transformation statistical effects

Since measures of time to complete a task are theoretically lognormally distributed, log transformations of the raw data were examined. However, since the data contains values of zero (0), logs cannot be taken directly from the raw data, since the log of zero (0) cannot be computed. One recommendation is to add a constant of one (1) to each score.[8] However, in several cases of performance measures, the raw data is actually categorized continuous data where all orders, for example, completed in the same day as initiated were assigned a zero. In these cases there are no "true" zero values since each order takes some time to complete. The lower bound of each interval is taken as the performance result, leaving the lowest interval with a value of zero.

Suggesting that some value in the middle of the interval defined by each integer may be a more appropriate representation of the interval, staff asked Dr. Gleason to determine the optimal constant for the transformation. Dr. Gleason simulated continuous lognormal distributions which he then categorized using the performance data categories. Using a MathCad© worksheet (Attachment 3 to this appendix), Dr. Gleason then

---

[8] *Id*.

calculated a constant that, when added to the actual categorized data, would best represent the parameters of the original continuous distribution. Staff used Dr. Gleason's worksheet to calculate the constant that would best fit actual ILEC and CLEC performance.[9] The worksheet calculates that the upper limit for the optimal constant is approximately 0.5, with virtually all values between 0.3 and 0.5.[10] The mean and standard deviation of the distributions affect the value of the constant, making the optimal constant slightly different for each analysis.

These results are theoretically reasonable as well. The mathematical midpoint of an interval in a skewed distribution[11] is not likely to accurately represent the distribution of scores in most intervals.

The sixteen distributions were then log-transformed using three different constants: 0.3, 0.4, and 0.5. The preponderance of transformations resulted in a closer approximation to normality. The results are presented in Attachment 4. In a few of the small samples where transformations did not improve normality, the transformed results are still relatively close to normality. From these transformations it is difficult to tell which constant best and most consistently improved normality. The transformations made with the constant of 0.5 improved normality most for large means, and the transformation with the constant of 0.3 improved normality most for the samples with small means.

---

[9] Staff used the worksheet by entering an actual posted result with a constant added to the mean. The standard deviation was entered as posted since it does not change when a constant is added to the data. If the added constant matched the calculated estimated constant (designated "α"), the calculated constant was taken to be the optimal constant. For example, for an actual posted performance result with a mean of 0.92 and a standard deviation of 3.46, a mean of 1.28 (0.92 + 0.36) was entered as the final "guess" confirming that the optimal constant would be 0.36.

[10] A survey of average-based results for January through June, 2000, indicates that the theoretically most appropriate constant ranges between 0.3 and 0.5 for about 99 percent of the results (687 out of 696). For about one percent of the results (7 out of 696) a constant of between .06 and 1.2 appeared to be most appropriate. However, these seven results occurred only in March 2000 and only once for each of seven submeasures, and the constants for all other months were 0.5 for each of these submeasures. Constants for three other submeasure results were estimated to be about 0.25, but for each submeasure, the other months had estimated constants of over 0.3. No results indicate optimal constants less than 0.3 for any result since March 2000. Constants under 0.3 appear to be anomalies.

[11] E.g., the interval 0.0 to 1.0 has a mathematical midpoint of 0.5.

Since it would be impractical to calculate the optimal constant for each result, and one of the constants must be selected, staff examined the sensitivity of the modified *t*-test to the different transformations. Staff sought to determine which one of the constants would result in the least discrepancy compared to using an optimal constant for each result. Type I error probabilities (α) were calculated for the sixteen submeasures. Attachment 5 presents a comparison of *t*-test results using raw scores and different log transformations. Use of the 0.4 constant for all results appears to minimize potential discrepancies between using result-specific optimal constants versus using a single constant for all results. In other words, compared to using the 0.3 or 0.5 constants for all results, using the 0.4 constant results in an α closer to the α calculated by using the result-specific optimal constant.

The constant should be added wherever the average-based performance measures produce zeros by categorizing continuous data. Following this criterion, current information indicates that constants should be used with transformations for performance measures 7, 14, 21, 28, and 37, and that performance measures 1 and 44 need no constants.

The constant should be added at the level of categorization. For example, if the smallest measurement unit is one day, then a constant of 0.4 days should be added to each observation. If the smallest measurement unit is one-hundredth of an hour, such as for performance measure 21, then the constant of 0.4 of one-hundredth (4 thousandths) of an hour should be added to each observation.

Performance evaluation implications

The meaning of the impact on actual performance result decisions was also examined. (See Attachment 5.) Fourteen of the 16 industry-aggregate results had the same “pass/failure” designation for the log transformation analysis as they did for the original raw score analysis. Two results failed the log-based analysis when they originally passed the raw score analysis: submeasure examples “9” and “10.” Both showed CLEC average performance worse than ILEC average performance, but the average differences were not statistically different using a raw score based test. Submeasure example “9” illustrates the nature of this failure. The original data analysis resulted in a “pass” with an alpha of 0.32, whereas the log transformed data resulted in a “failure” with an alpha of less than 0.0001. For both submeasure examples “9” and “10,” compared to the means, the medians showed greater differences between ILEC and CLEC performance. In these instances, the log transformation has the effect of giving a better reflection of the difference between the two distributions than was the case for the raw data. In the raw data analyses, extreme scores in the ILEC distribution mask the typically poorer performance for the CLEC. The transformation minimizes the effects of these extreme values.

Possibly the best illustration of the difference between the log and raw score based analyses is the results for one CLEC-specific result in submeasure example "9," where the direction of the difference between the medians was the opposite for the means. Whereas the raw score CLEC mean was 2.8 days, compared to the ILEC mean of 3.0 days, the CLEC median was 3.0 and the ILEC mean was 1.0. In other words, the average time to complete the OSS task for the CLEC was 2.8 days, which is better than the 3.0-day average for the ILEC. In contrast, the median for the ILEC performance was one day while the median for the CLEC performance was two days. In other words, the ILEC took one day or less to complete the OSS task for fifty percent of its customers, whereas the ILEC took three days or less to complete the same OSS task for fifty percent of the CLECs' customers. Similar to the submeasure "9" and "10" aggregate results, the median difference for this CLEC result was much larger than the mean difference. These results show that the distributions are markedly different. Whereas the raw score analysis did not show this result to be significant ($\alpha = 0.86$), the transformation analysis identified a significant difference ($\alpha < 0.0001$).[12]

In these cases, the log transformation analysis appears to track the differences in medians more closely than the means. This is consistent with academic sources that point out the value of median-based assessments when the data is skewed. For example, Hays (1997) states,

---

[12] A reverse of this situation is demonstrated in the BANY hypothetical data set provided in Verizon's comments in John Jackson's paper (see References). Results for the BANY data set show performance "failure" for the raw score analysis, but performance "success" for the log transformed analysis. In the BANY data set, using a 0.10 critical significance level, the difference between the ILEC and CLEC means is significant for a raw score analysis ($p = 0.083$), and marginally significant for a permutation analysis (p = 0.0855 to 0.1085). However, one CLEC outlier is responsible for the CLEC mean being greater than the ILEC mean. With this outlier of 53.0, the CLEC mean is 9.9 compared to the ILEC mean of 8.3. However, without the outlier, the CLEC mean is less than the ILEC mean, 5.6. The cumulative distribution for this hypothetical data is presented in Attachment 6 and illustrates that typically the CLEC received better (simulated) performance than the ILEC. In this case, the log transformation analysis reflects typical performance more than average performance in that reduces the effect of the outlier and identifies these results as a performance "success" (p = 0.97).

> The alteration of the score for a single extreme case in a distribution may have a profound effect on the mean. It is evident that the mean follows the skewed tail in the distribution, but the median does so to a lesser extent. The occurrence of even a few very high or very low cases can seriously distort the impression of the distribution given by the mean, provided that one mistakenly interprets the mean as the typical value. If you are dealing with a nonsymmetric distribution and you want to communicate the typical value, you must report the median. (p. 181)

To assess the operational meaning of the transformation staff also examined the cumulative distribution. The cumulative percentage distribution data and graphs are included in Attachment 1 for submeasures 1 through 16, and in Attachment 7 for the CLEC-specific results described above. While a frequency distribution shows the number of "orders" completed for each time interval, a cumulative distribution shows the percentage of "orders" that were completed by a specific number of "days" or less.[13] For example, the frequency distribution for submeasure "9" shows that approximately 500 of the ILEC's orders took 3 days to complete. In contrast, the cumulative distribution for submeasure "9" shows that about 80 percent of the ILEC's orders were completed in three days *or less*. In the cumulative distribution graphs, the higher line represents better service. For example, for submeasure "9," ILEC customers (black line) are getting better service than CLEC customers (white line) up until the point where approximately 80 percent of the orders have been completed – at a time interval of "3." Where the ILEC line is higher means that compared to CLEC customers, a greater percentage of ILEC customers are getting their orders completed within the specified time intervals. After this point, CLEC customers are getting better service as illustrated by the lines crossing. After this point the CLEC line is higher, meaning that a higher percentage of CLEC customers are getting their "orders" completed within the same time interval as ILEC customers.

[13] The terms "days" and "orders" are used for illustrative purposes and do not necessarily represent the actual units. Because of possible proprietary data issues the actual terms are not used.

The graphs in these two attachments provide a more detailed comparison of the two distributions and illustrate what the log transformation analysis detects that the raw score analysis does not. The graph shows that for submeasure example “9,” for up to 80 percent of the customers, the ILEC gave better performance to its own customers. Specifically, 52 percent of ILEC customers’ orders, compared to 6 percent of CLEC customers’ orders, were completed within one day after the order was confirmed. Similarly, 72 percent of ILEC customers’ orders, compared to 31 percent of CLEC customers’ orders, were completed within two days after the order was confirmed. It is only in the final 20 percent of the distributions that CLEC customers received better performance than did the ILEC customers. For example, by the time five days passed, 88 percent of ILEC customers’ orders were completed compared to 95 percent of CLEC customer’s orders. These graphs depict distributions that are not “substantially equal,” where CLEC customers are predominately disadvantaged even though, on the average, their completion interval is less.

Conclusions

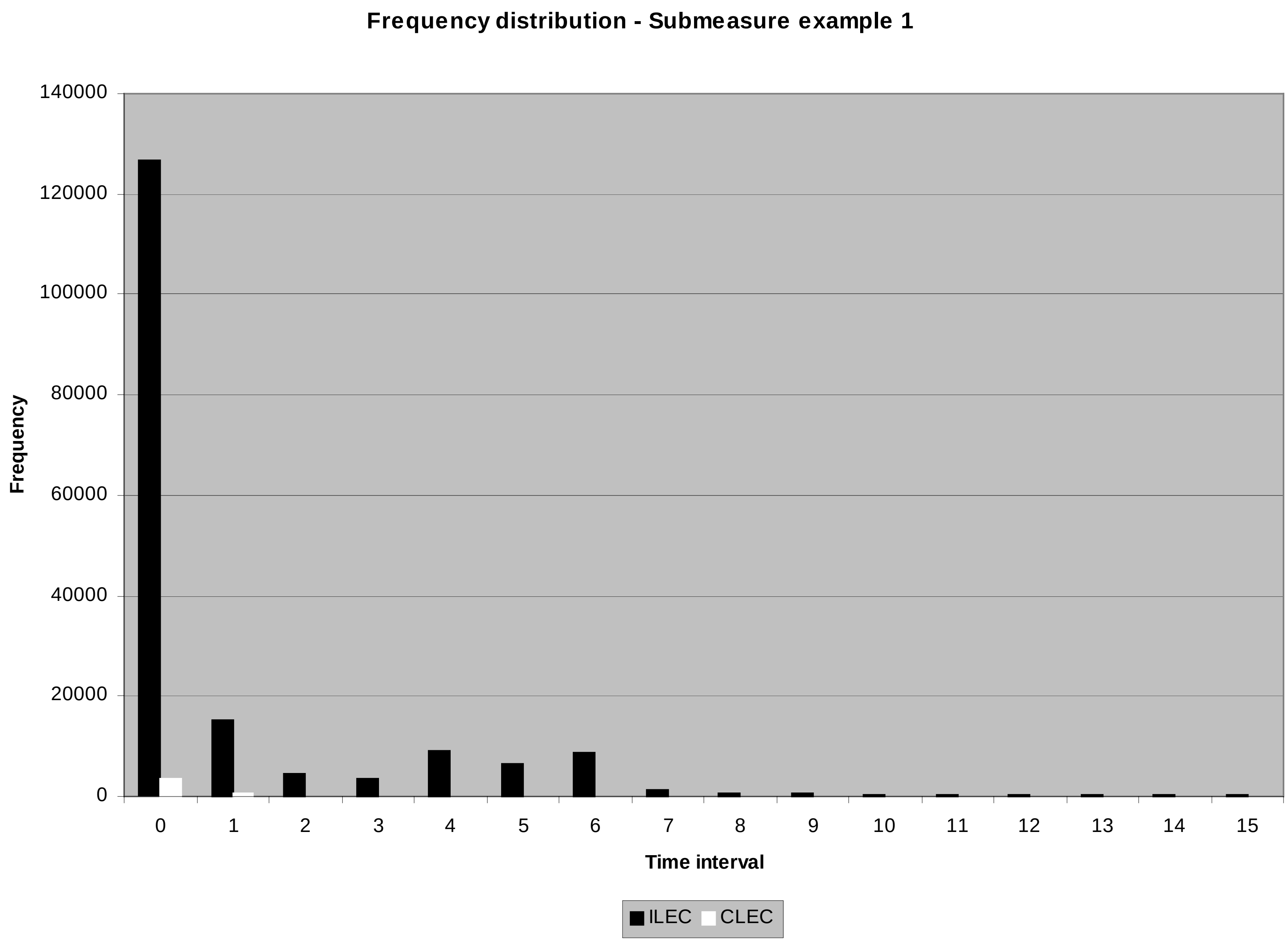
The log transformation of time measurement raw scores that have been increased by a constant is academically supported for average-based parity measures of time to complete a task. Using a constant of 0.4 (of the smallest interval) for each transformation reasonably corrects for distribution distortions introduced by categorizing all continuous values into integers, brings the data close to being normally distributed, and results in the least variation in a Type I error probability calculation compared to using a different constant for each transformation. Additionally, sample mean distribution normality is improved not only for small to moderate samples, but also for large samples. For these reasons, this log transformation of scores with an added constant of 0.4 is the best practical option for applying the modified *t*-test to average-based parity measures. Additionally, the log transformation allows an appropriate statistical and practically meaningful performance assessment. Until other methods are shown to be superior and ready to implement, the Modified *t*-test application using log transformations is justifiable and reasonable.

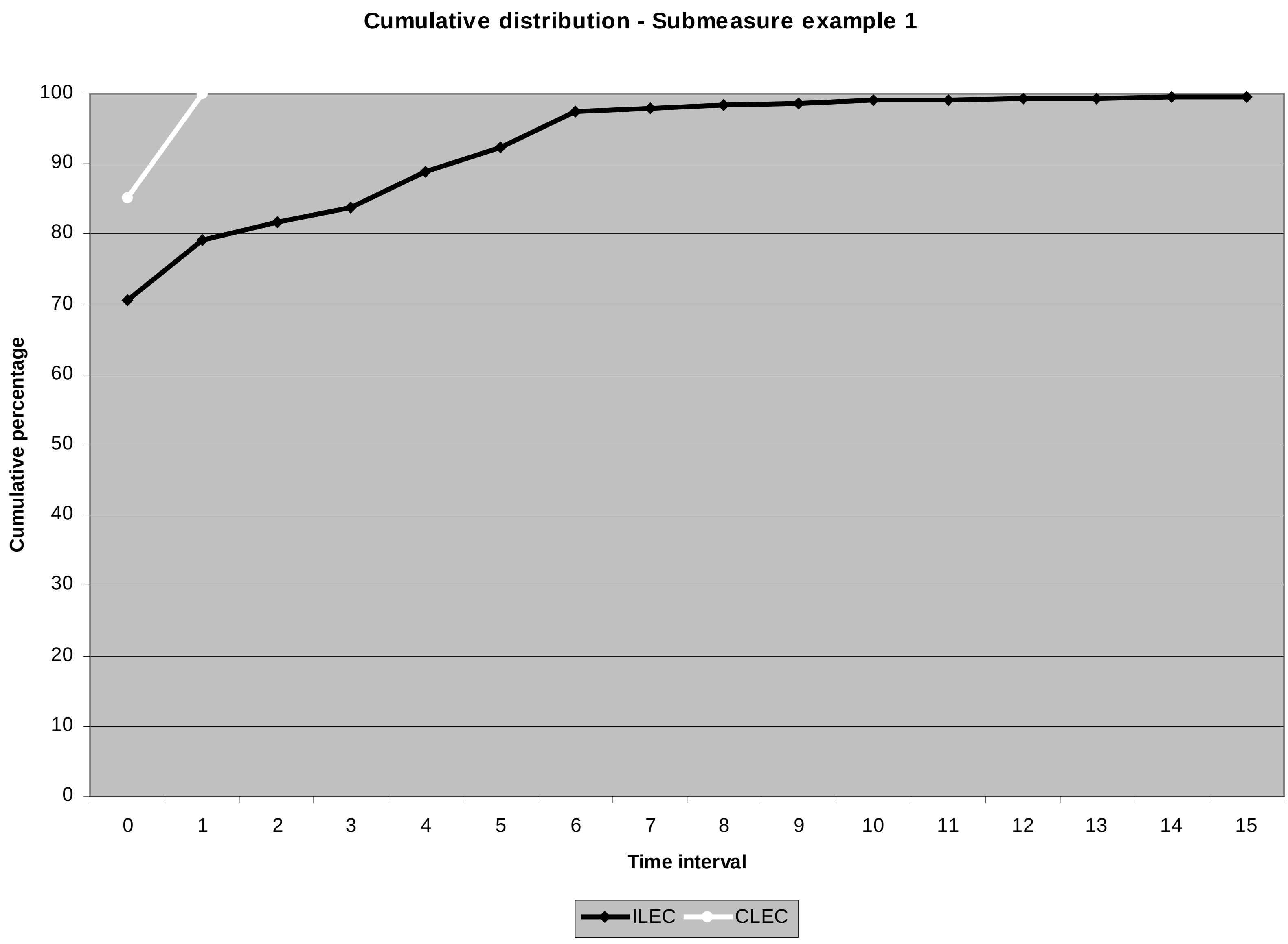
References

John D. Jackson, *Using permutation tests to evaluate the significance of CLEC vs. ILEC service quality differentials,* Verizon CA Opening Brief, Attachment 1 at Appendix 2 (April 28, 2000).

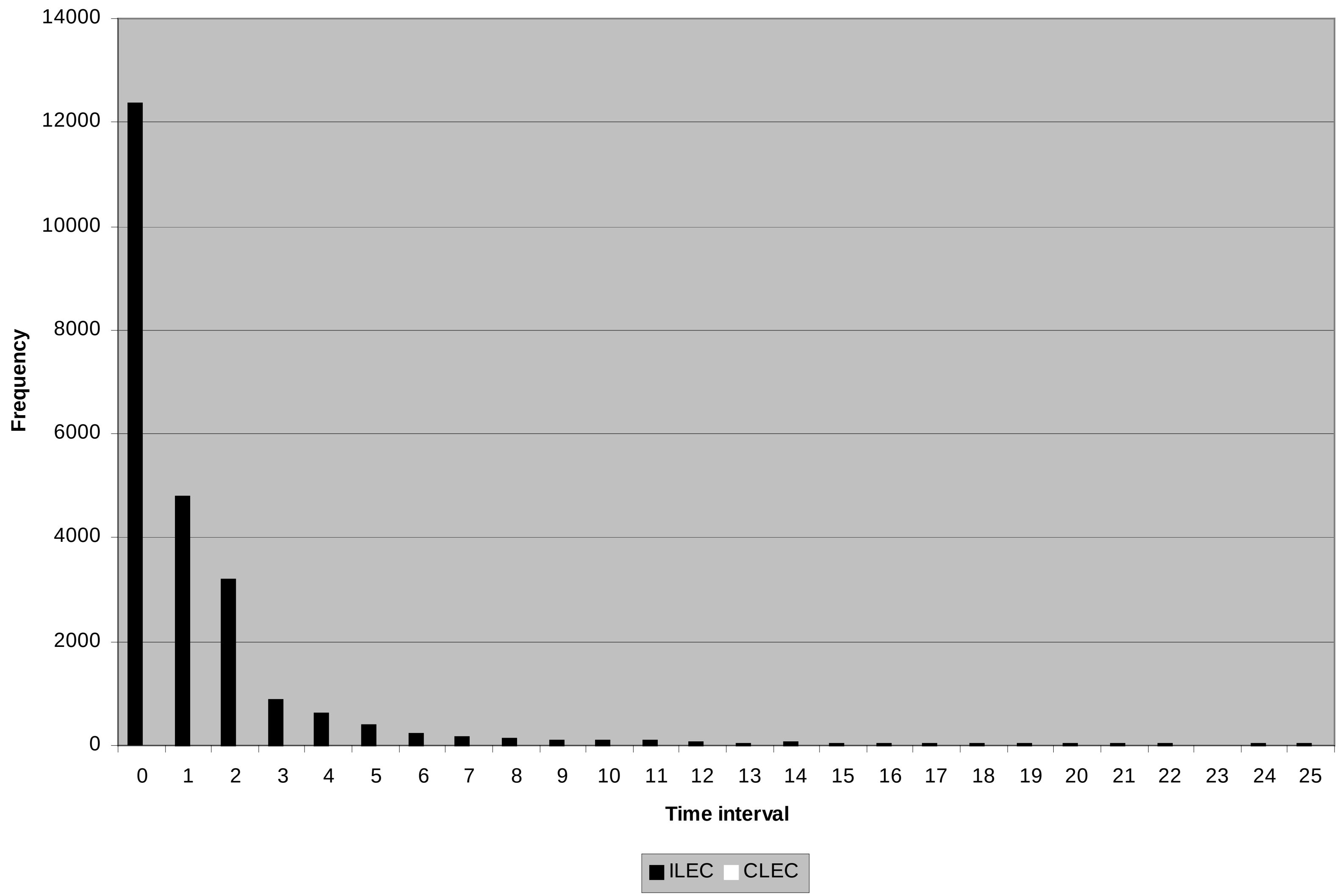
McNemar, Q. (1962). *Psychological statistics*. New York: John Wiley & Sons.

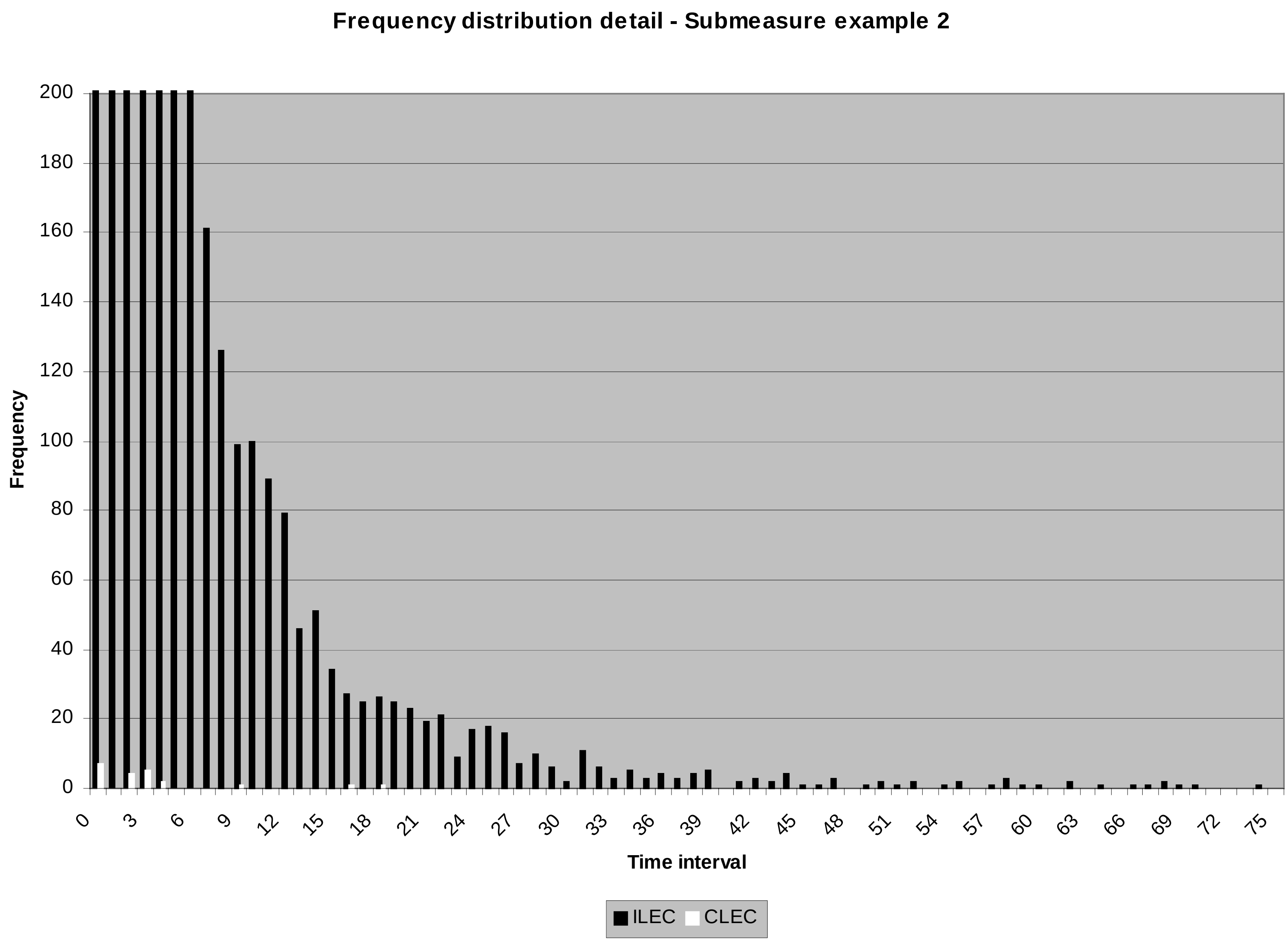
Winer, B.J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.


Frequency distribution - Submeasure example 1
140000
120000
100000
80000
60000
40000
20000
0
Frequency
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Time interval
ILEC
CLEC

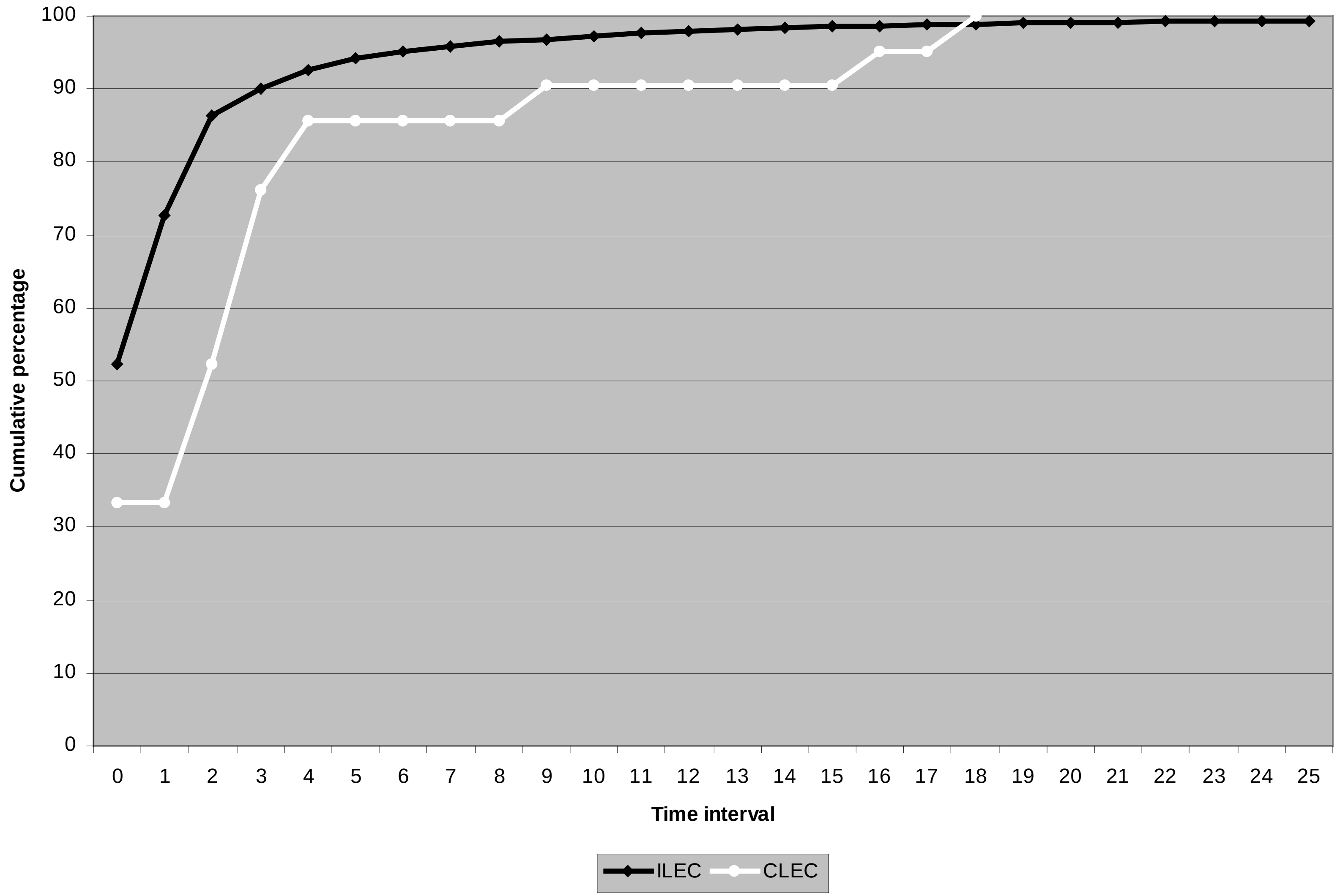

Cumulative distribution - Submeasure example 1
Cumulative percentage
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Time interval
ILEC
CLEC

Frequency distribution - Submeasure example 2

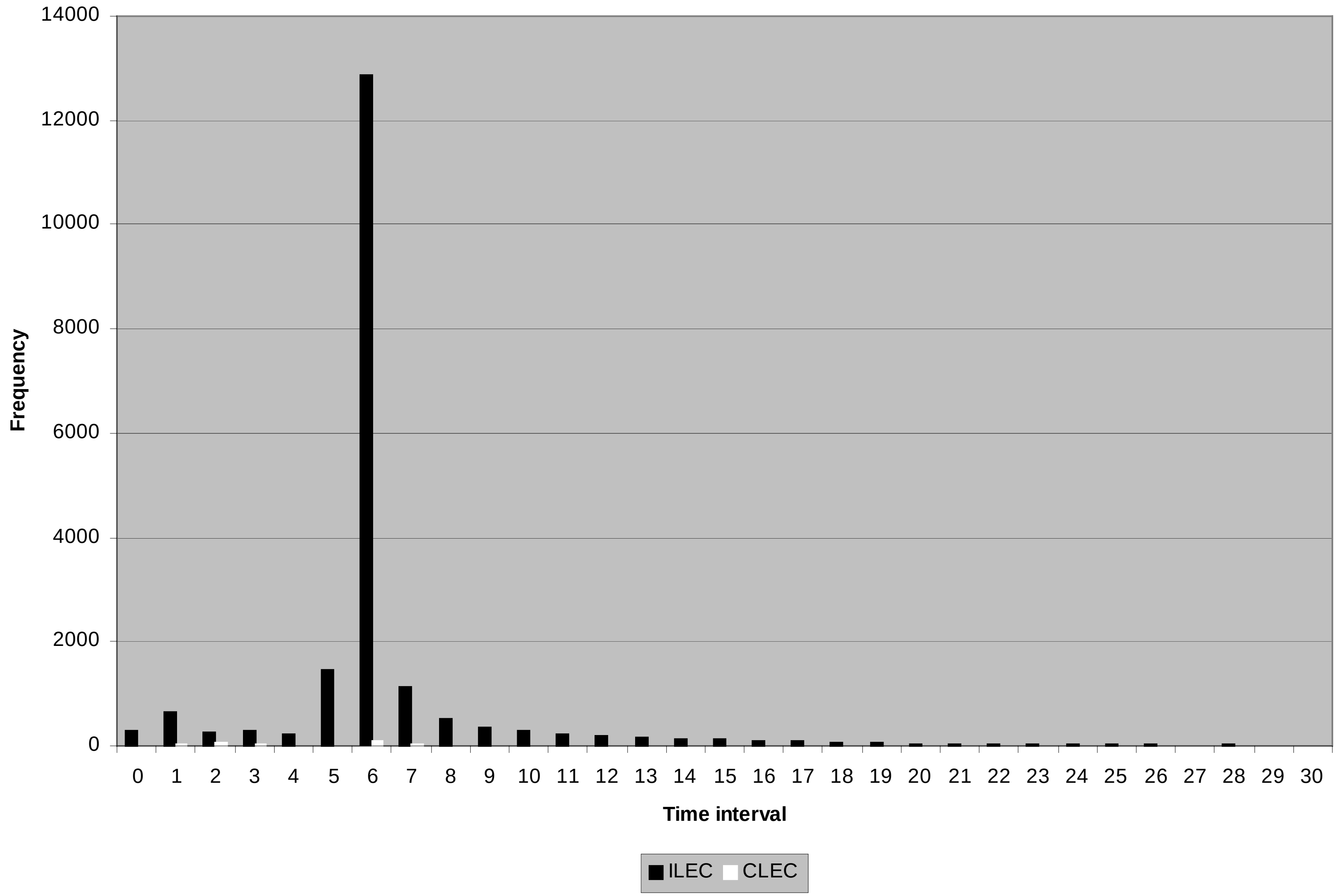


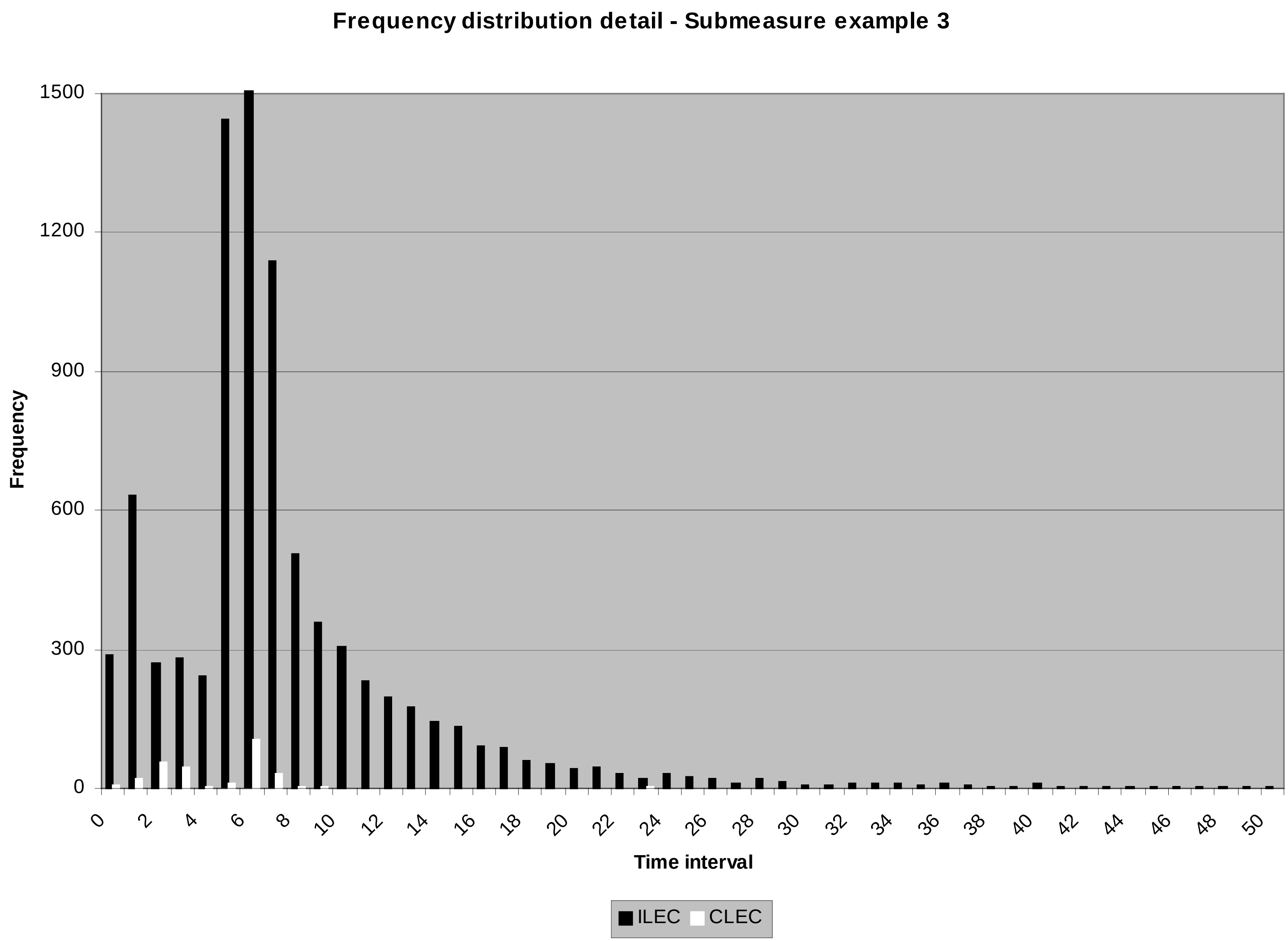

Frequency distribution detail - Submeasure example 2
Frequency
200
180
160
140
120
100
80
60
40
20
0
0
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
51
54
57
60
63
66
69
72
75
Time interval
ILEC
CLEC

Cumulative distribution - Submeasure example 2

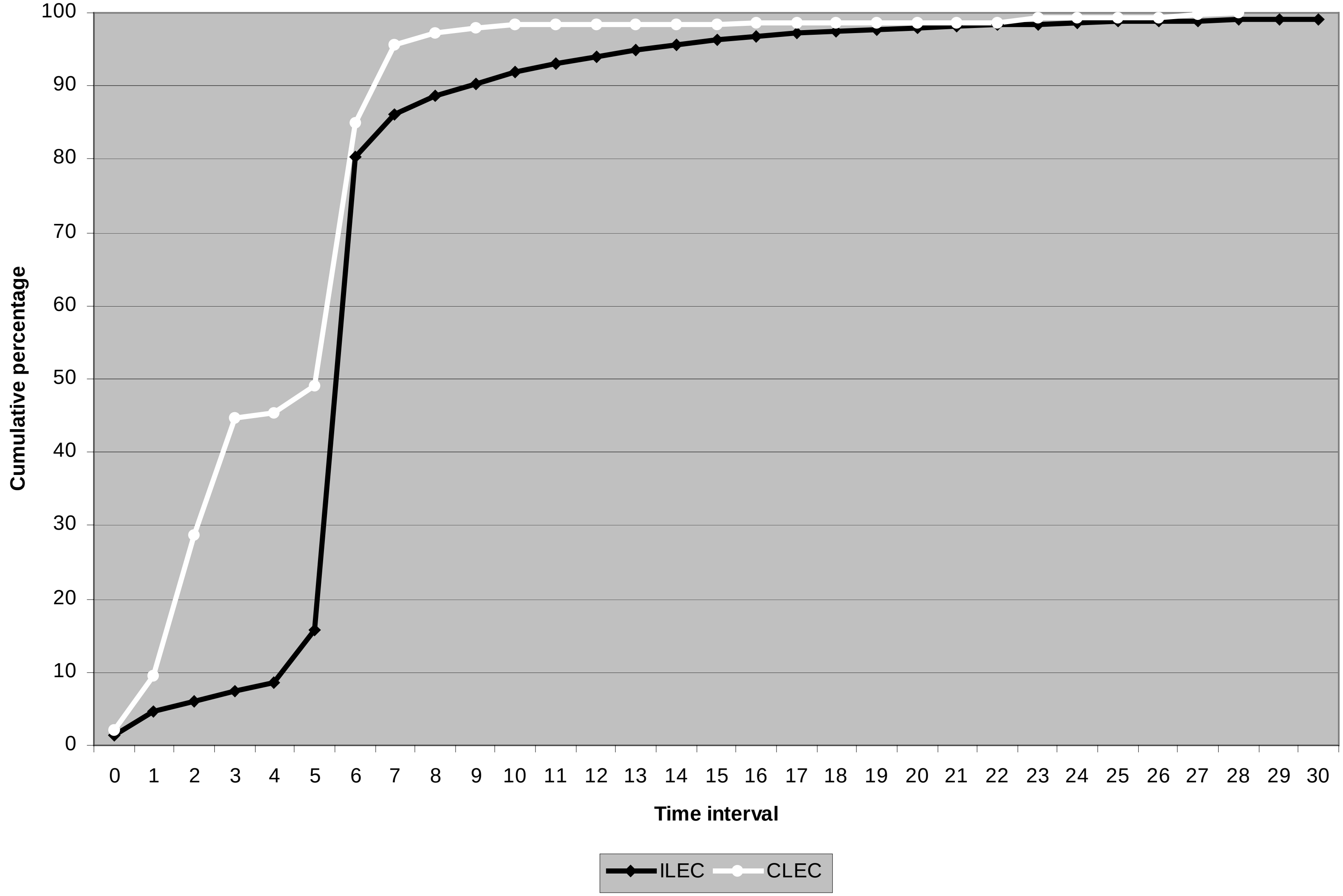


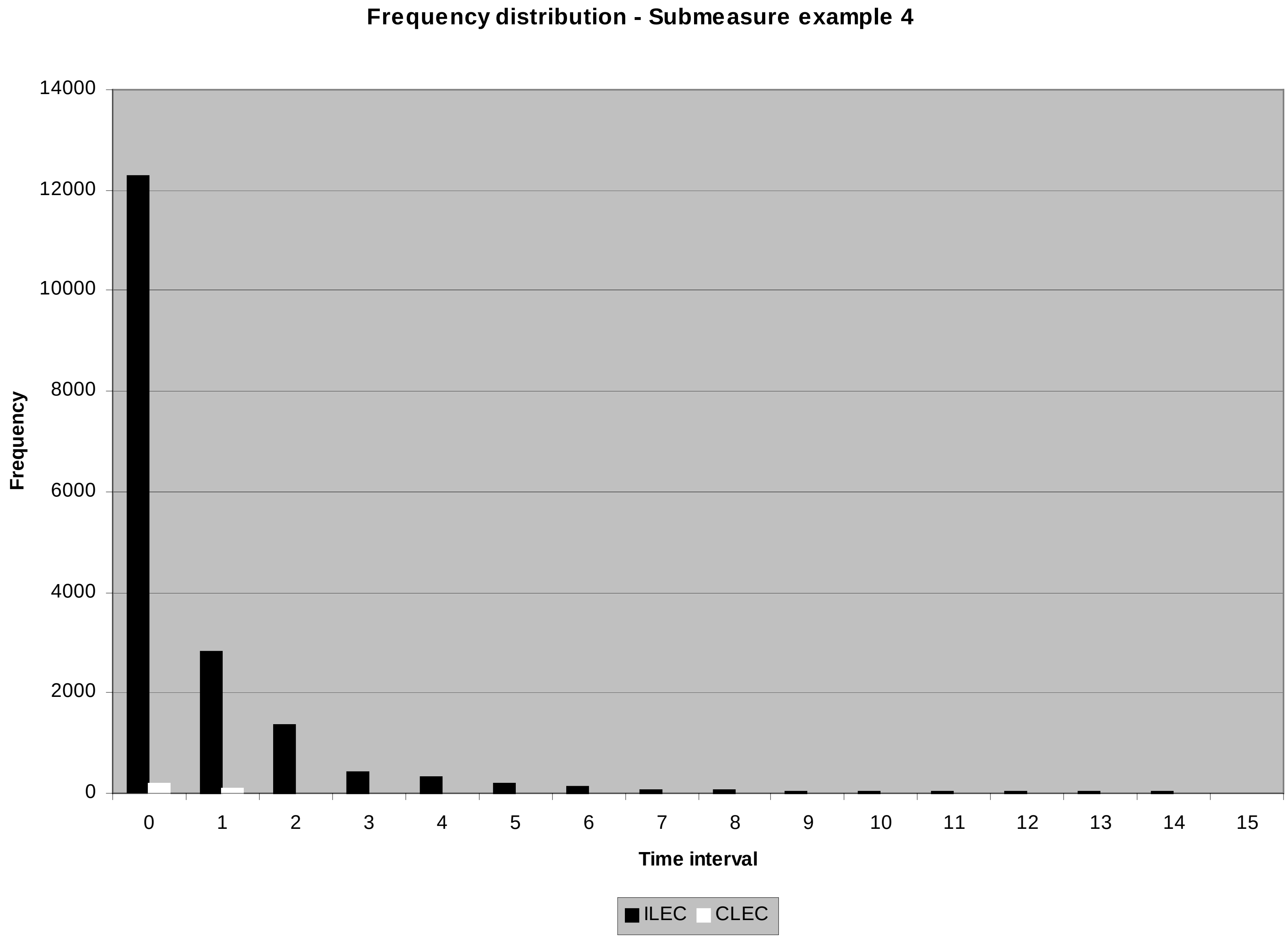
Frequency distribution - Submeasure example 3



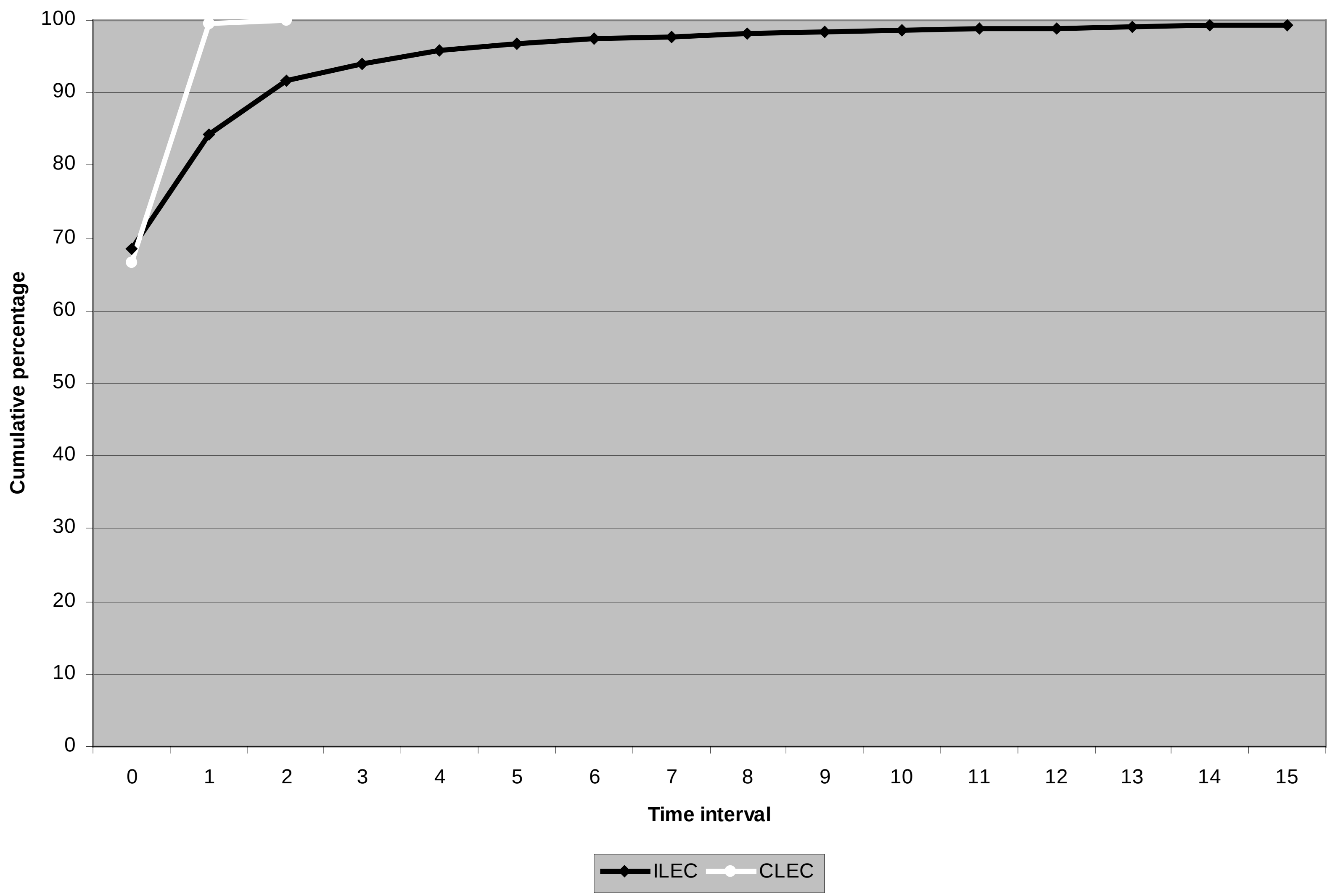

Frequency distribution detail - Submeasure example 3
Frequency
1500
1200
900
600
300
0
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
42
44
46
48
50
Time interval
ILEC
CLEC

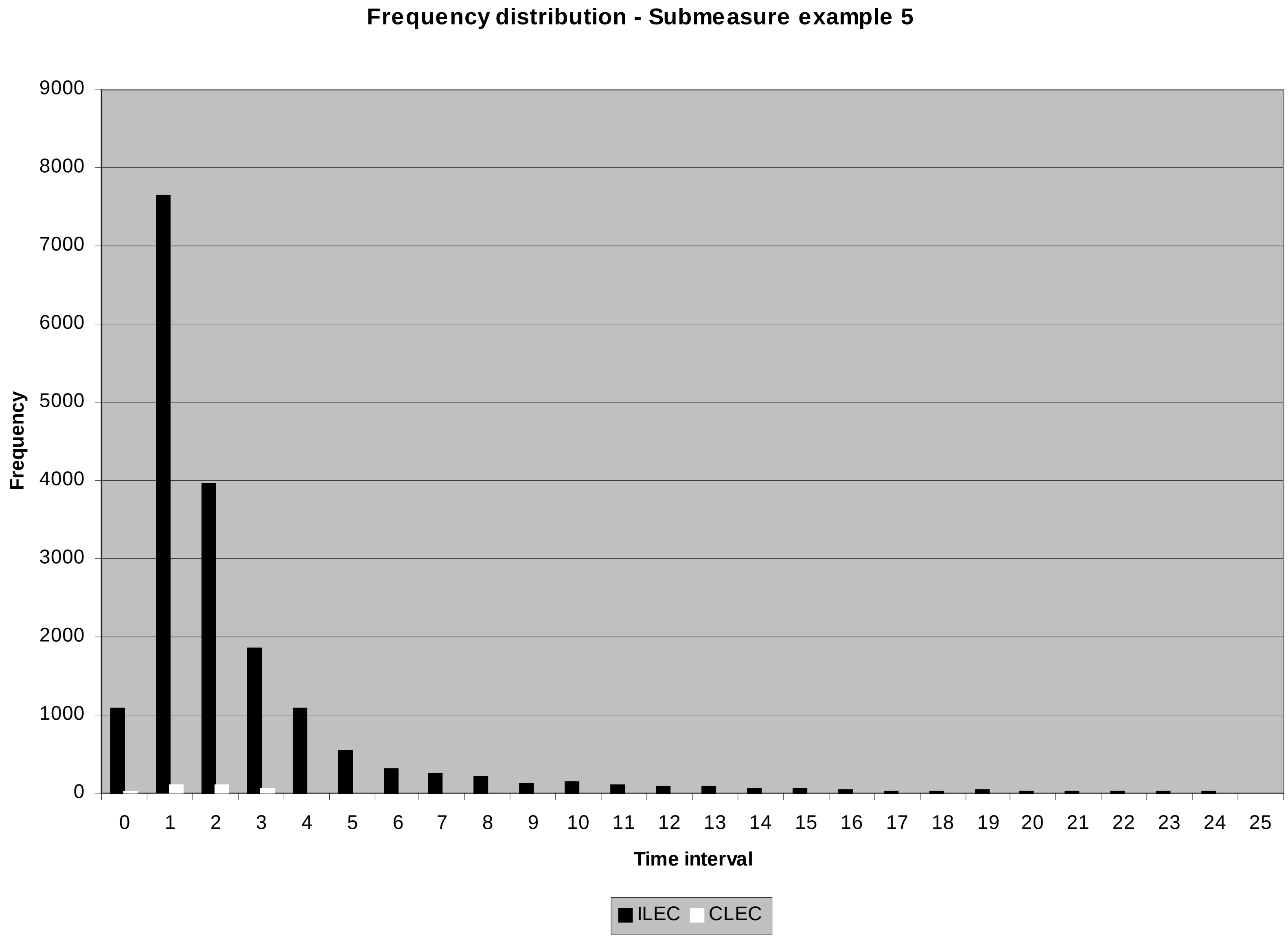
Cumulative distribution - Submeasure example 3




Frequency distribution - Submeasure example 4
14000
12000
10000
8000
6000
4000
2000
0
Frequency
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Time interval
ILEC CLEC

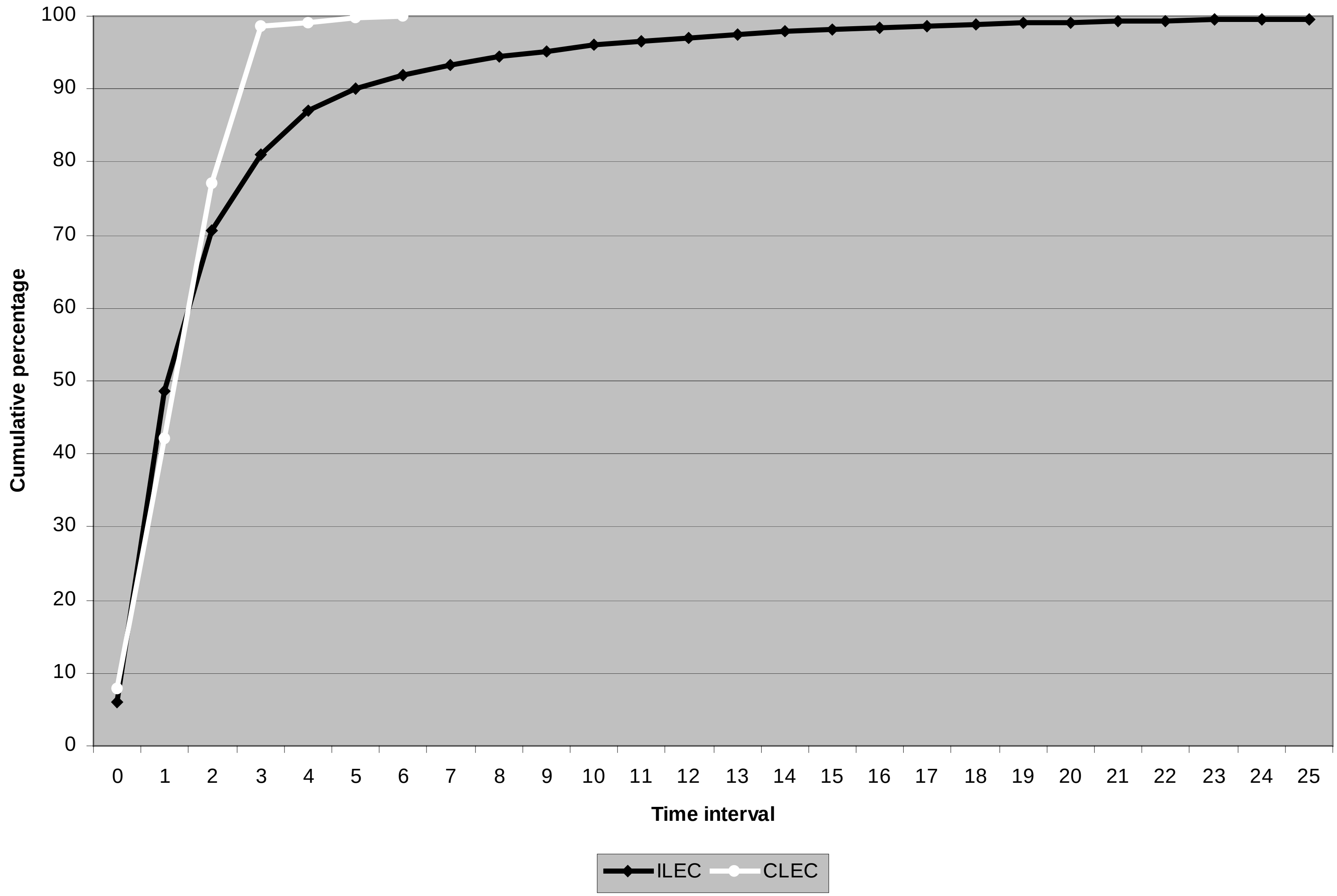
Cumulative distribution - Submeasure example 4

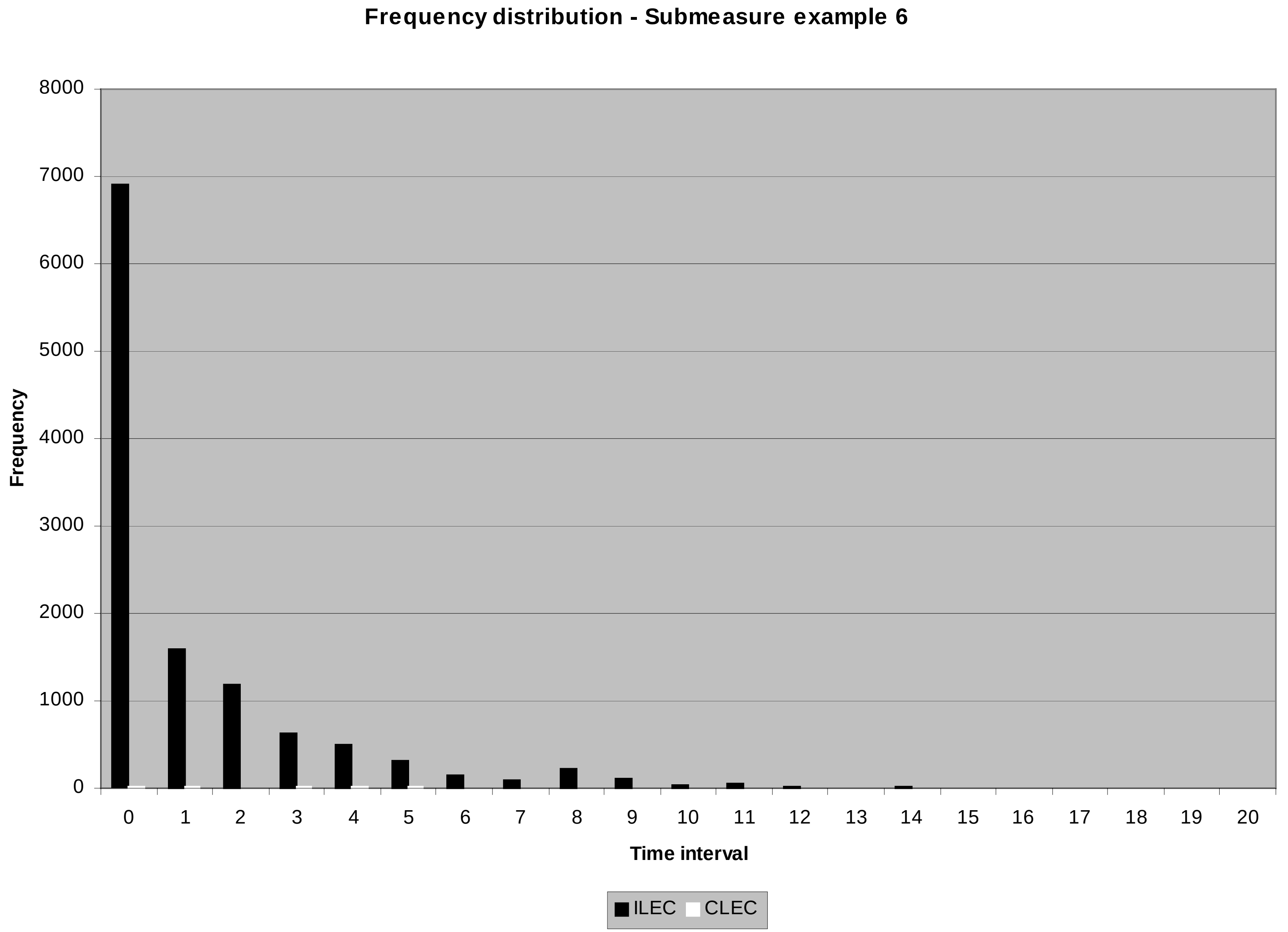


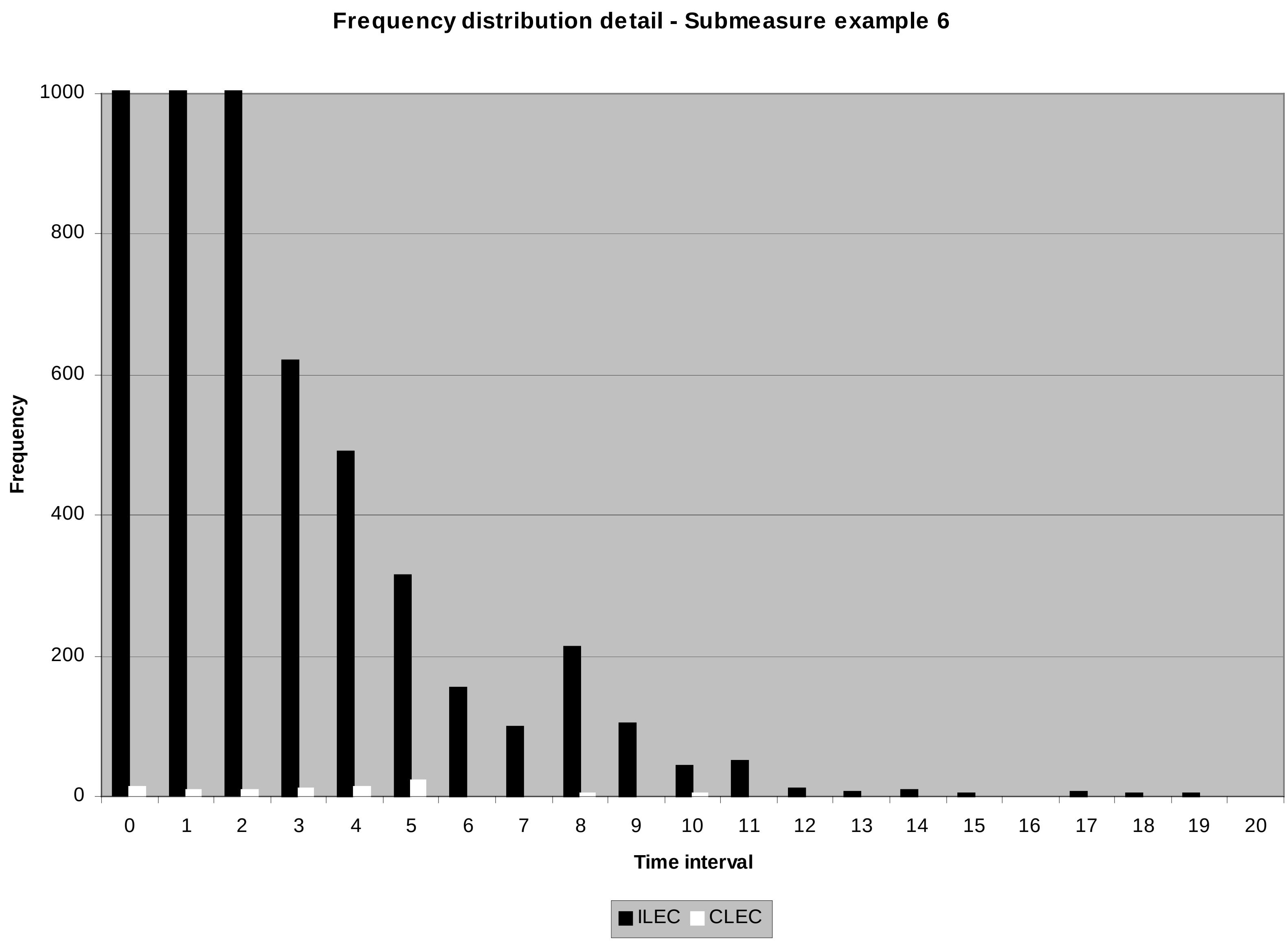


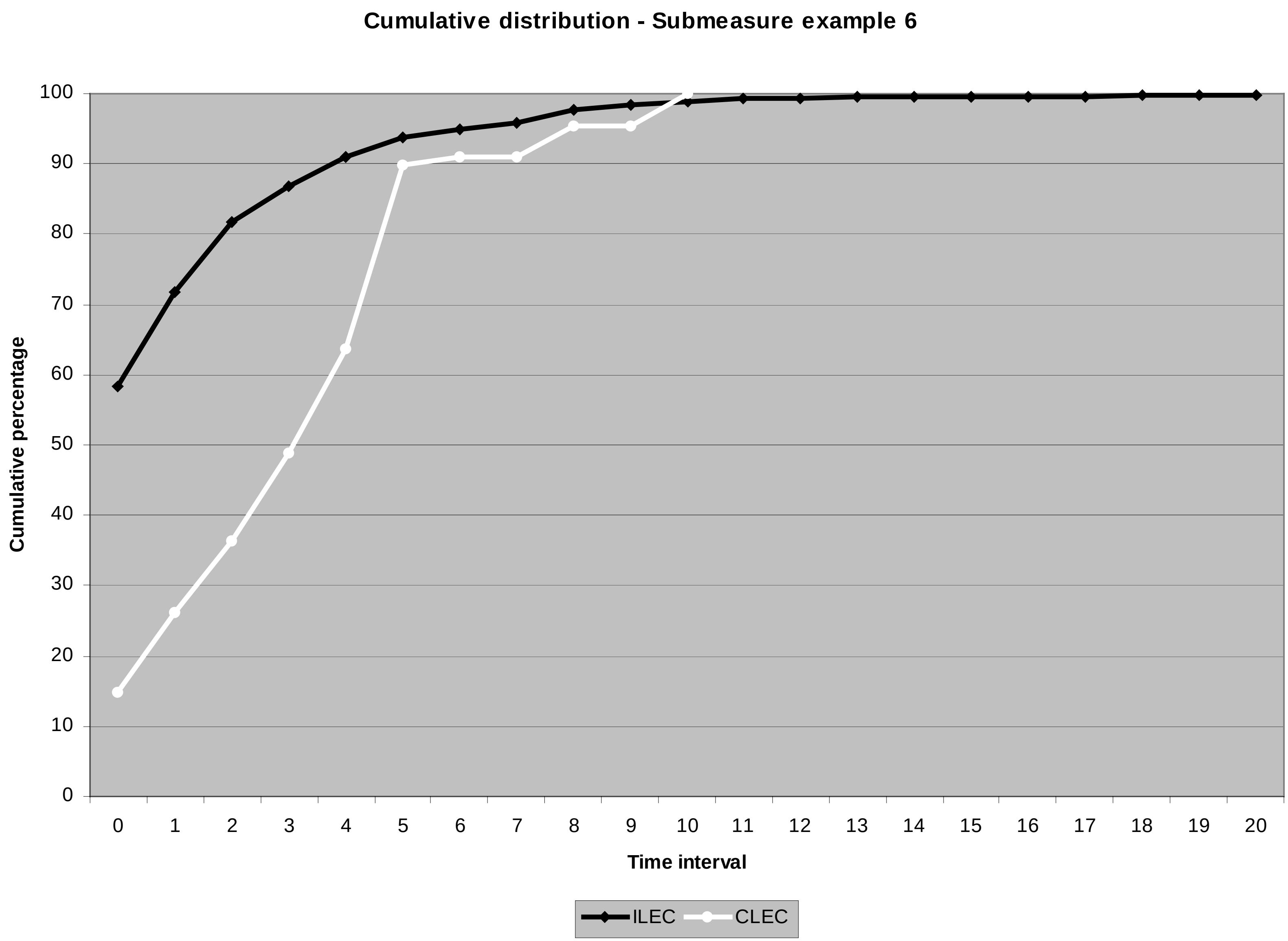
Frequency distribution - Submeasure example 5
Frequency
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
Time interval
ILEC
CLEC

Cumulative distribution - Submeasure example 5

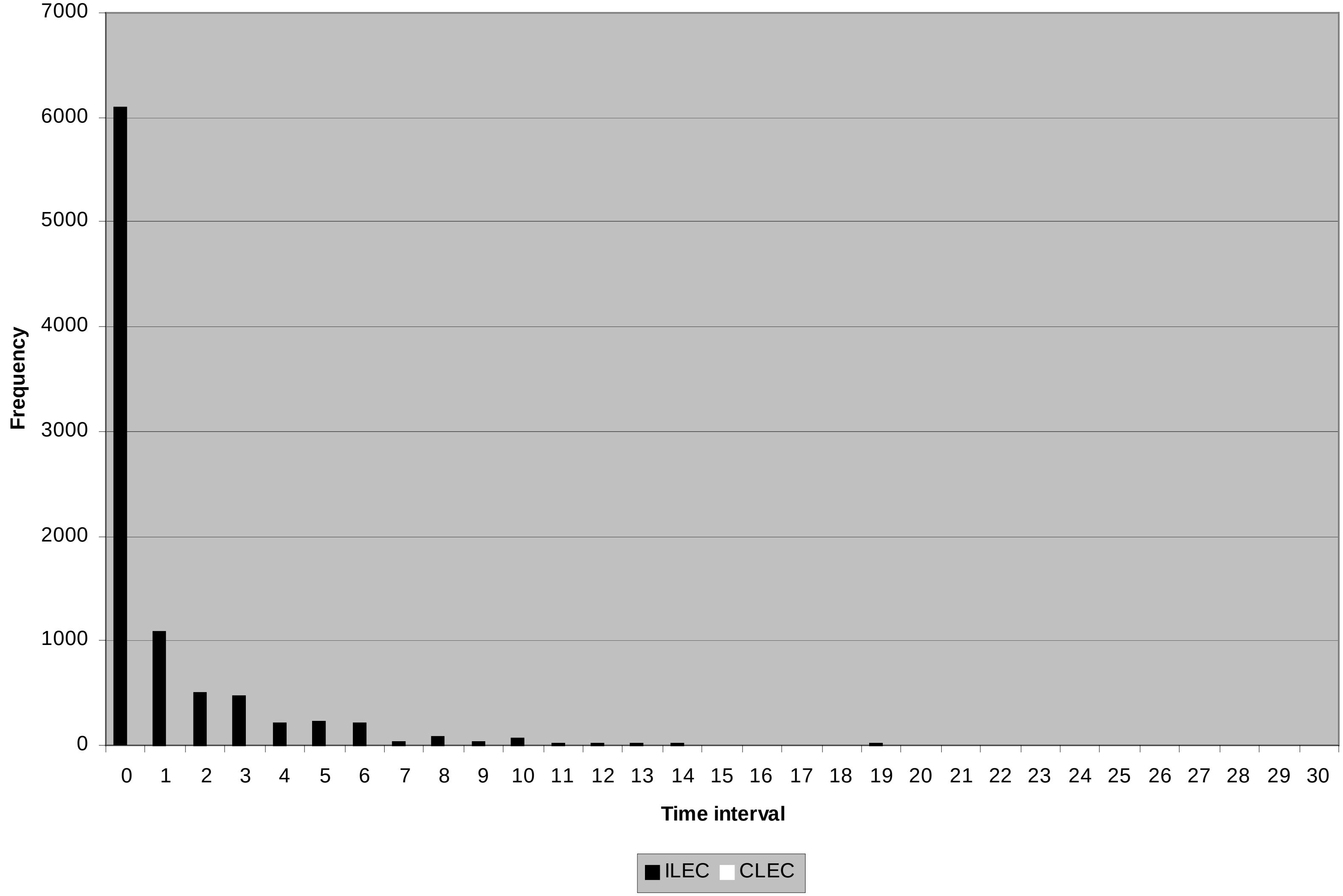



Frequency distribution - Submeasure example 6
Frequency
8000
7000
6000
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Time interval
ILEC
CLEC

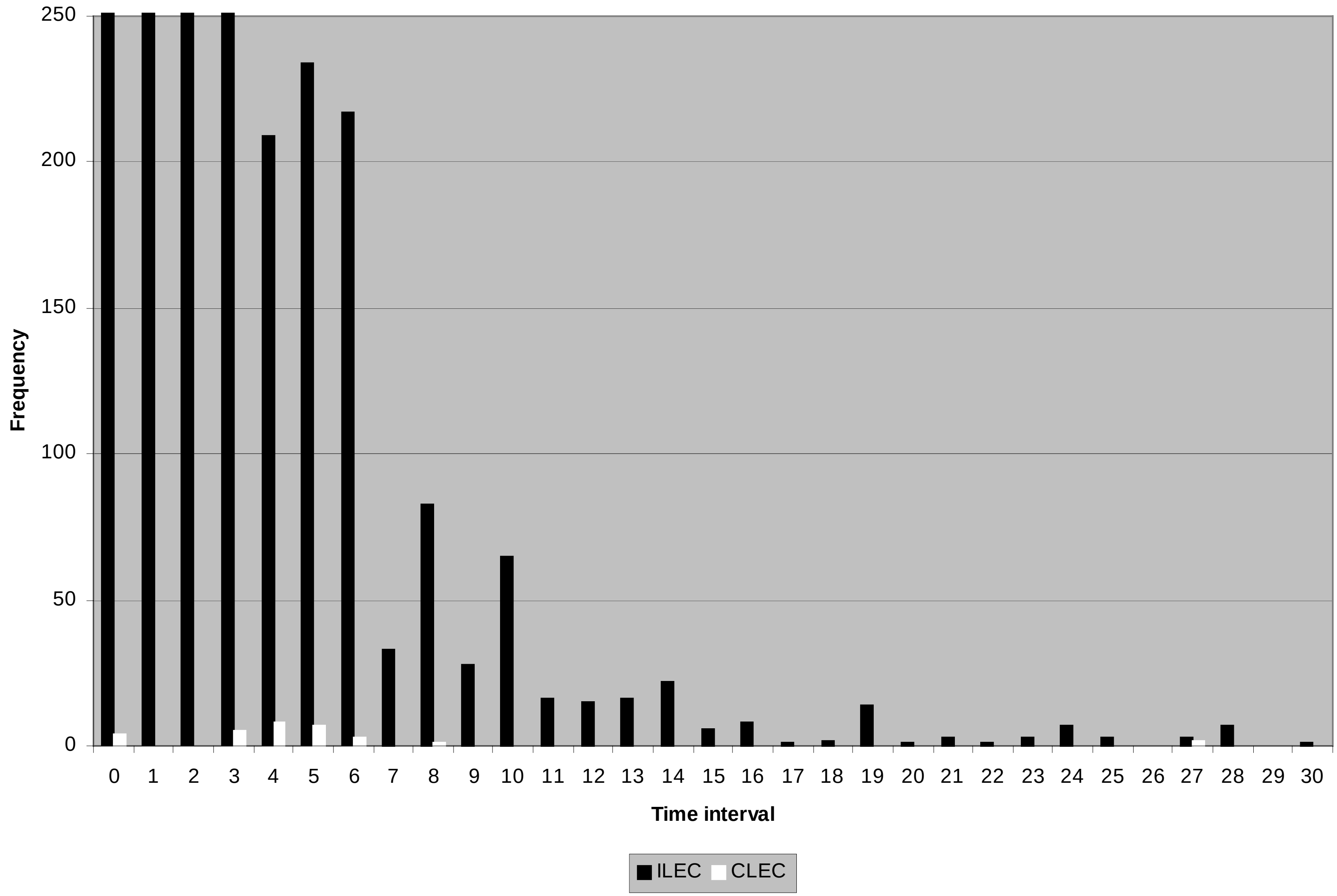

Frequency distribution detail - Submeasure example 6
Frequency
1000
800
600
400
200
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Time interval
ILEC
CLEC

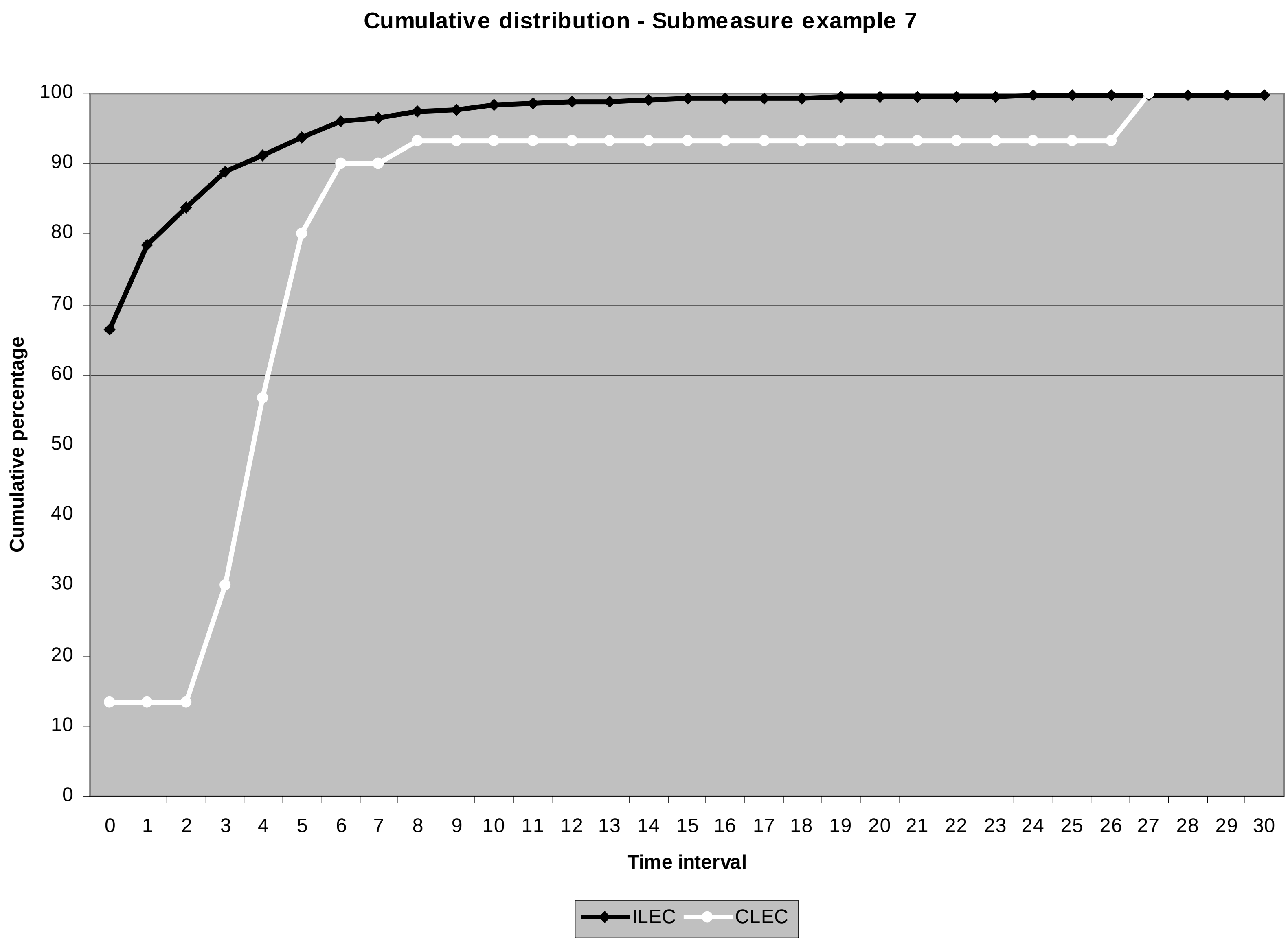

Cumulative distribution - Submeasure example 6
Cumulative percentage
100
90
80
70
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Time interval
ILEC
CLEC

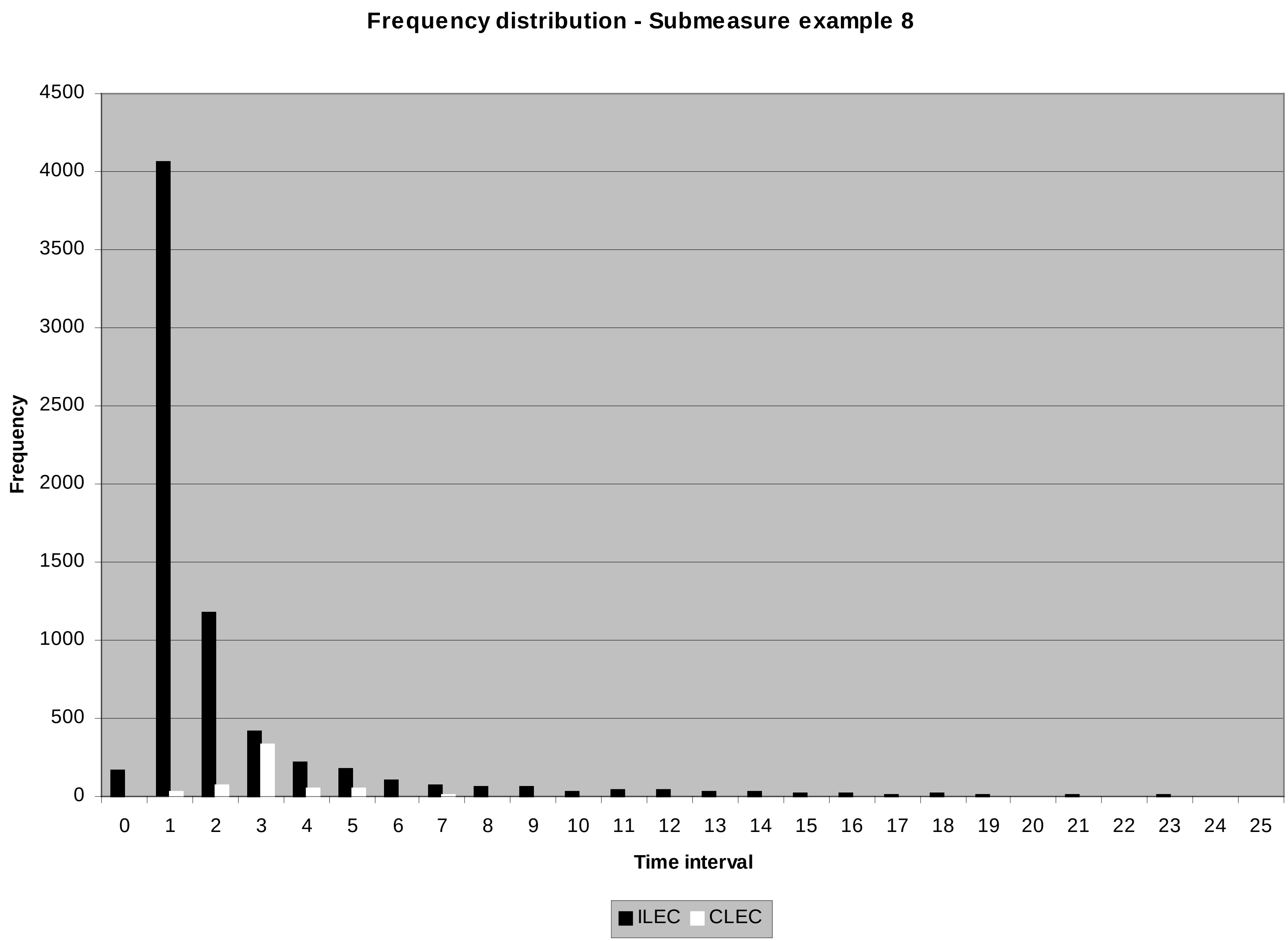
Frequency distribution - Submeasure example 7

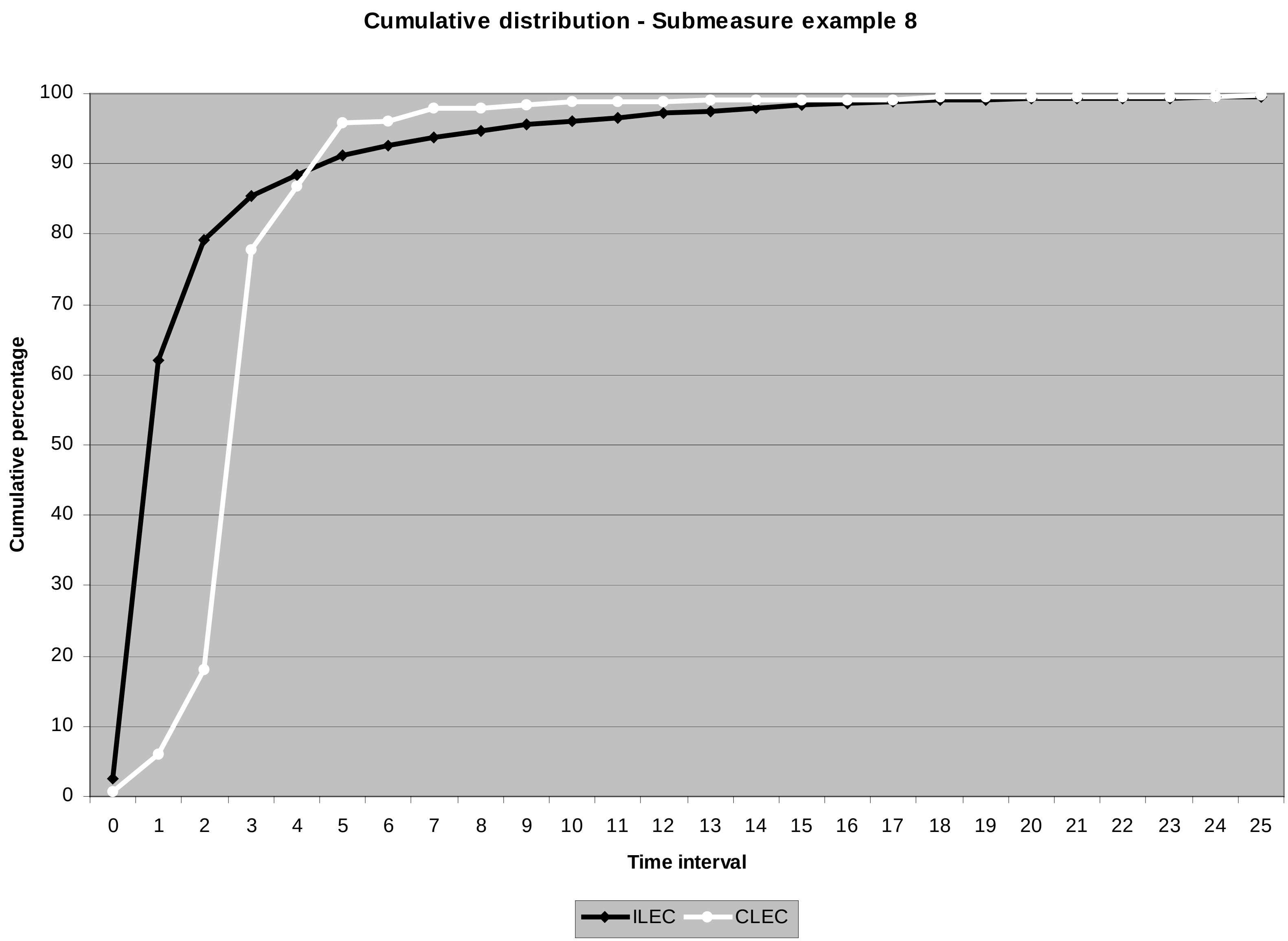


Frequency distribution detail - Submeasure example 7

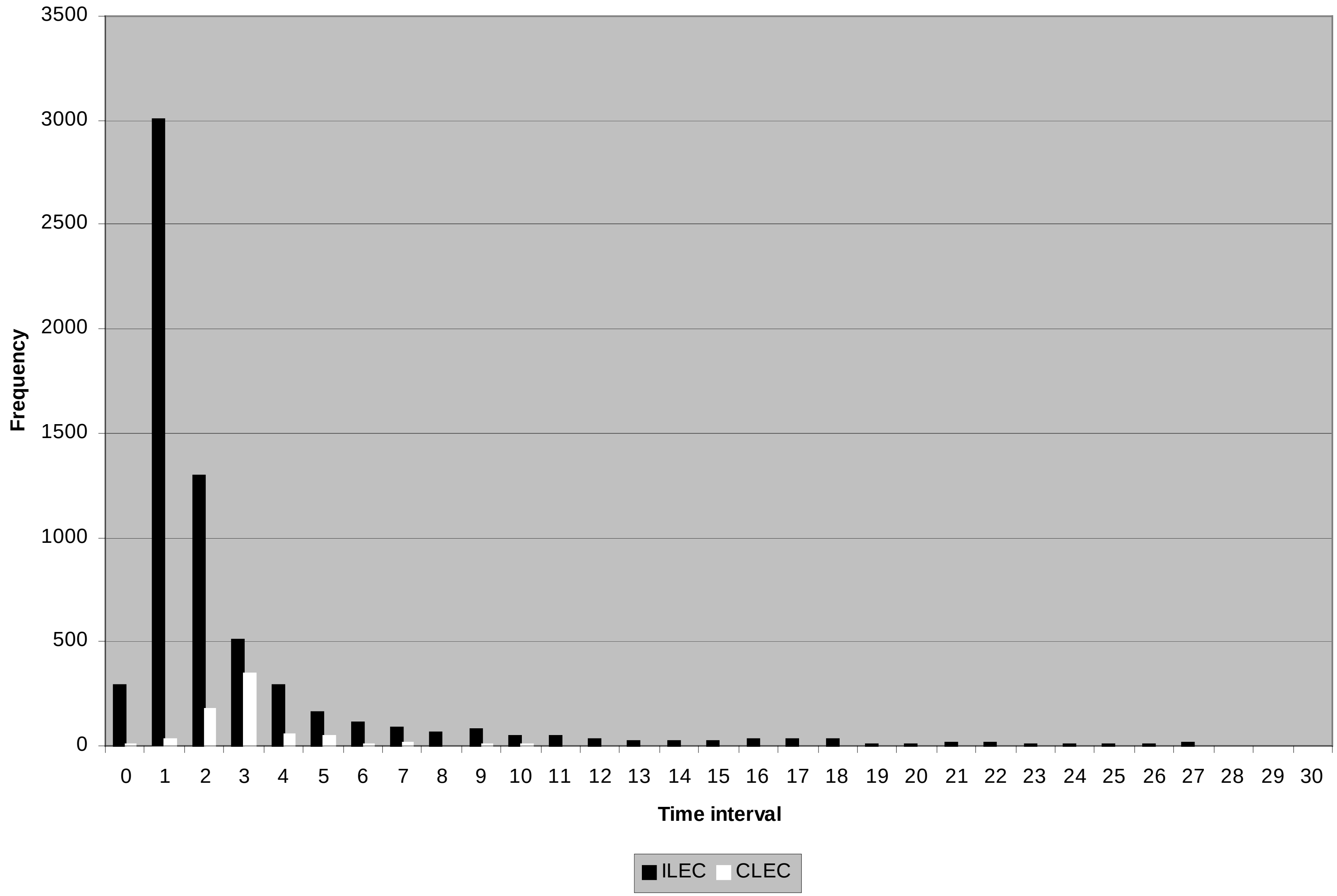


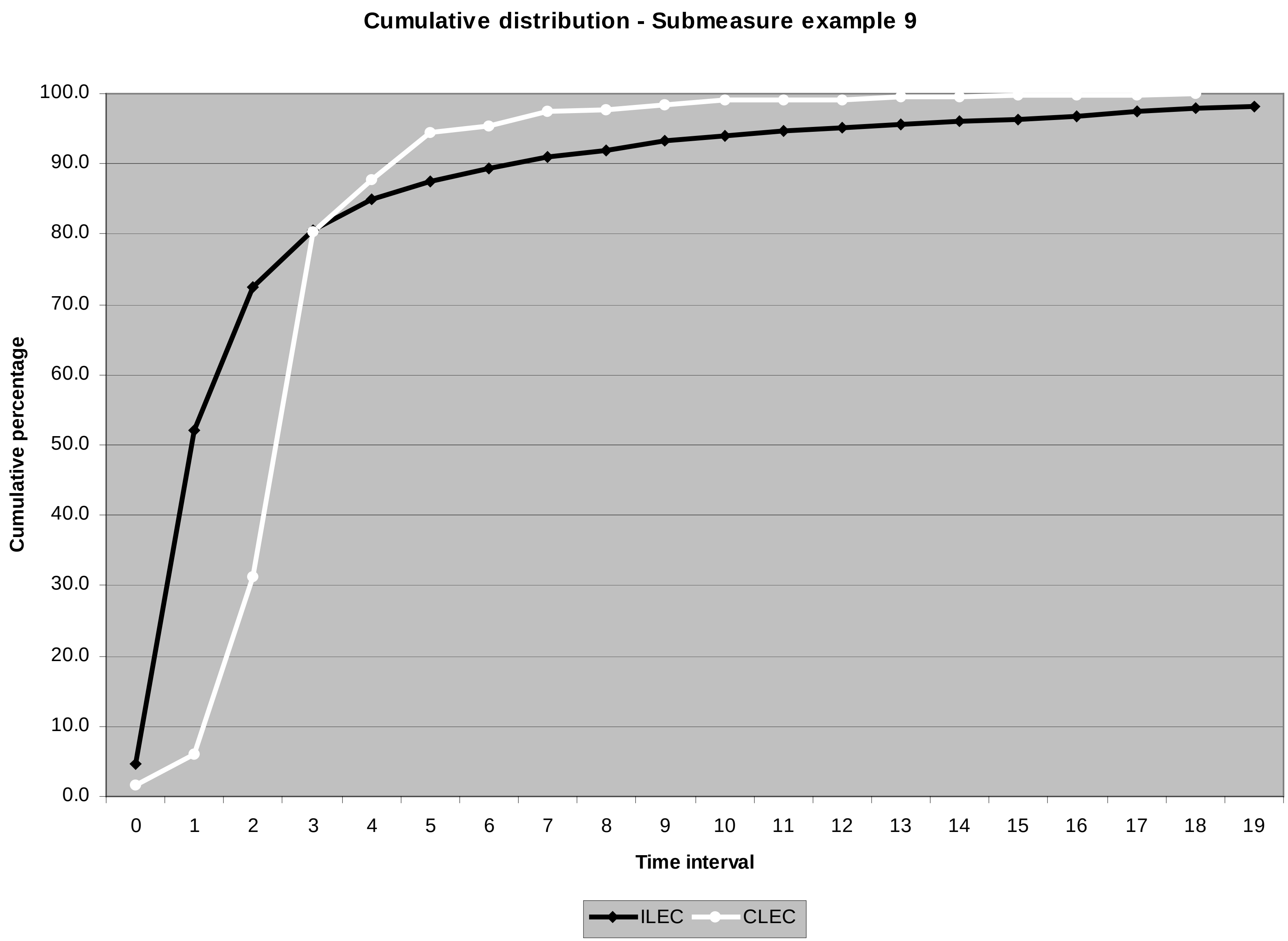

Cumulative distribution - Submeasure example 7
Cumulative percentage
100
90
80
70
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Time interval
ILEC
CLEC

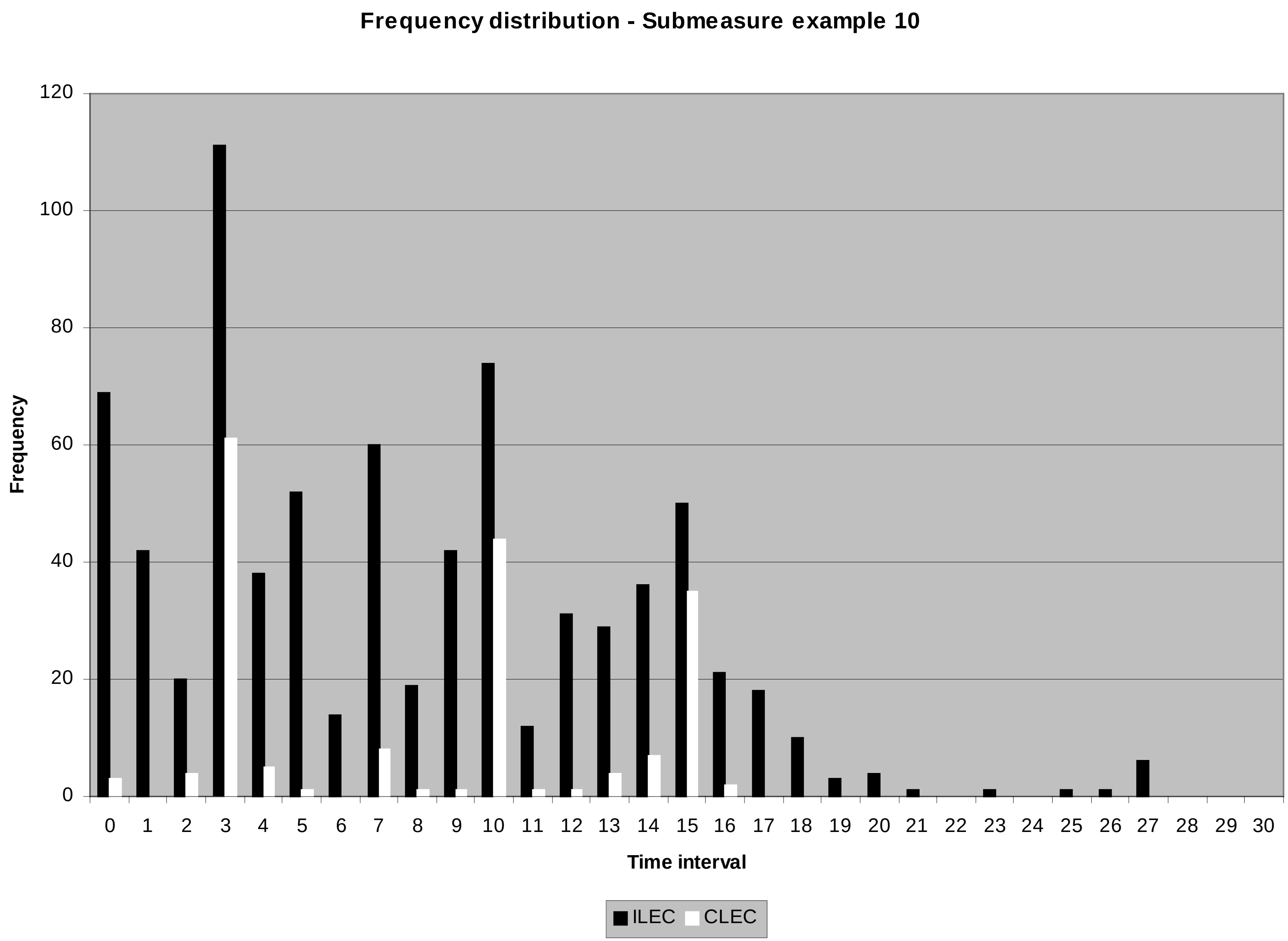

Frequency distribution - Submeasure example 8
Frequency
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Time interval
ILEC
CLEC

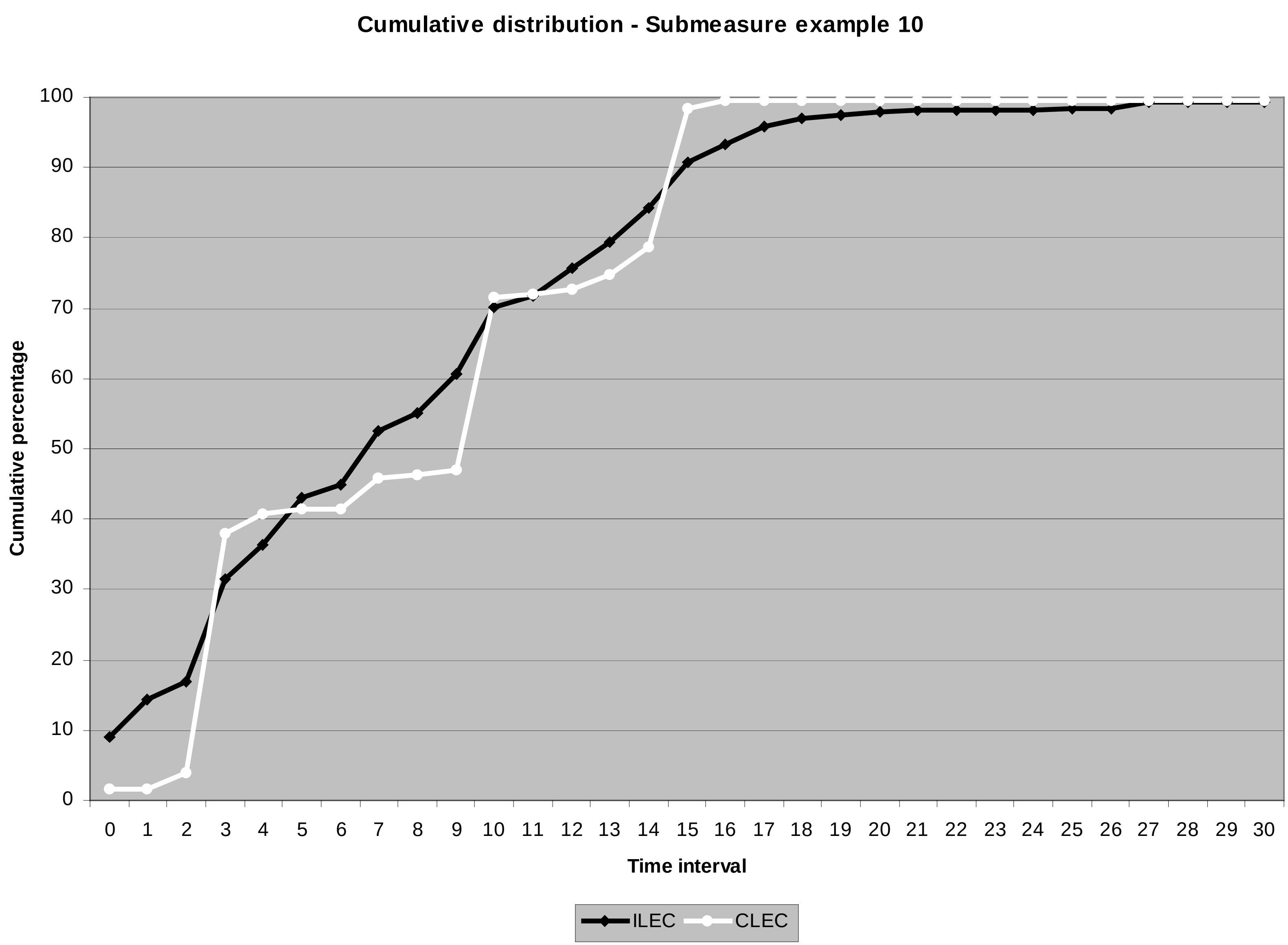

Cumulative distribution - Submeasure example 8
Cumulative percentage
100
90
80
70
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Time interval
ILEC
CLEC

Frequency distribution - Submeasure example 9

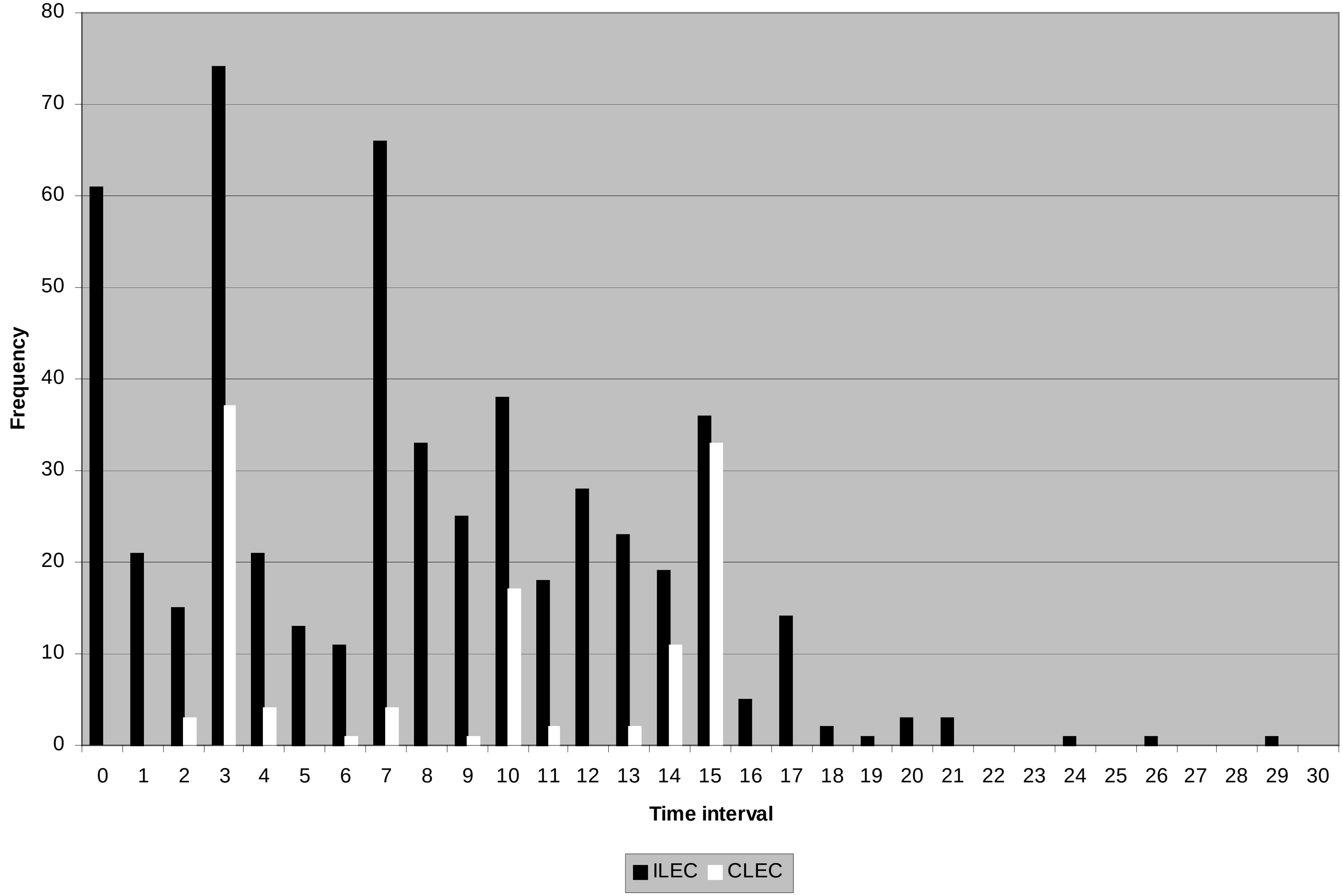



Cumulative distribution - Submeasure example 9
Cumulative percentage
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Time interval
ILEC
CLEC

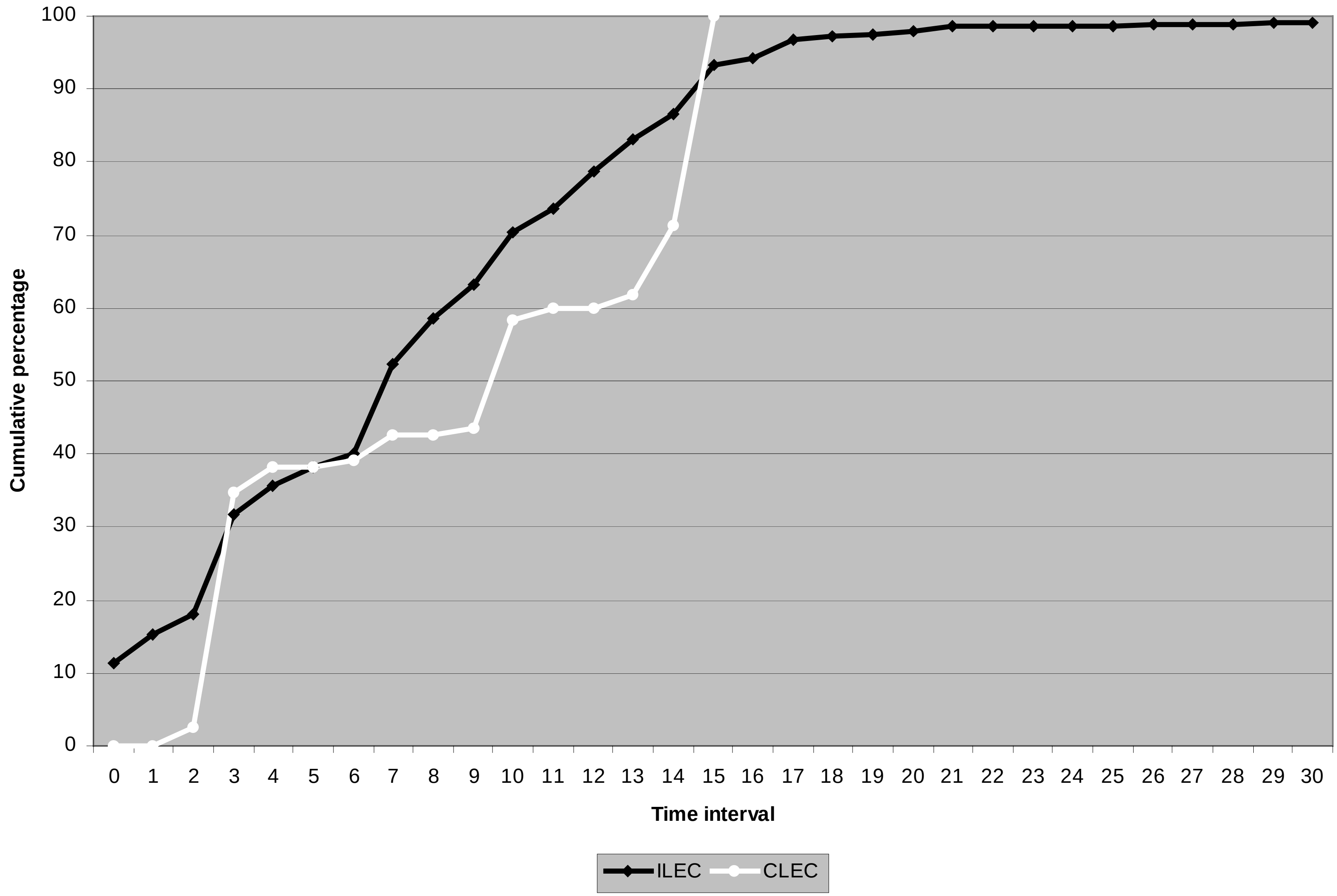

Frequency distribution - Submeasure example 10
Frequency
Time interval
ILEC
CLEC
0
20
40
60
80
100
120
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30


Cumulative distribution - Submeasure example 10
Cumulative percentage
0
10
20
30
40
50
60
70
80
90
100
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Time interval
ILEC
CLEC

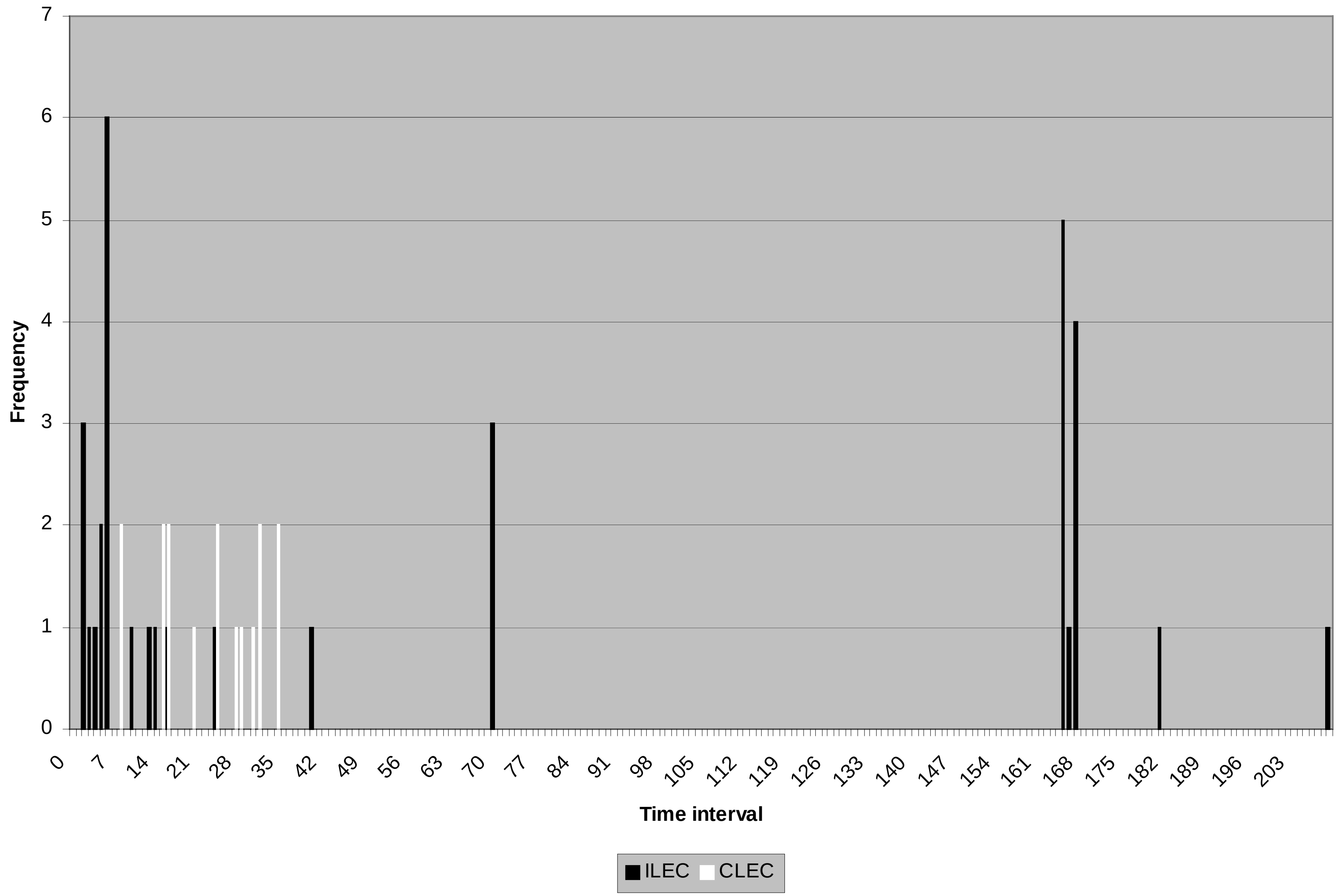
Frequency distribution - Submeasure example 11



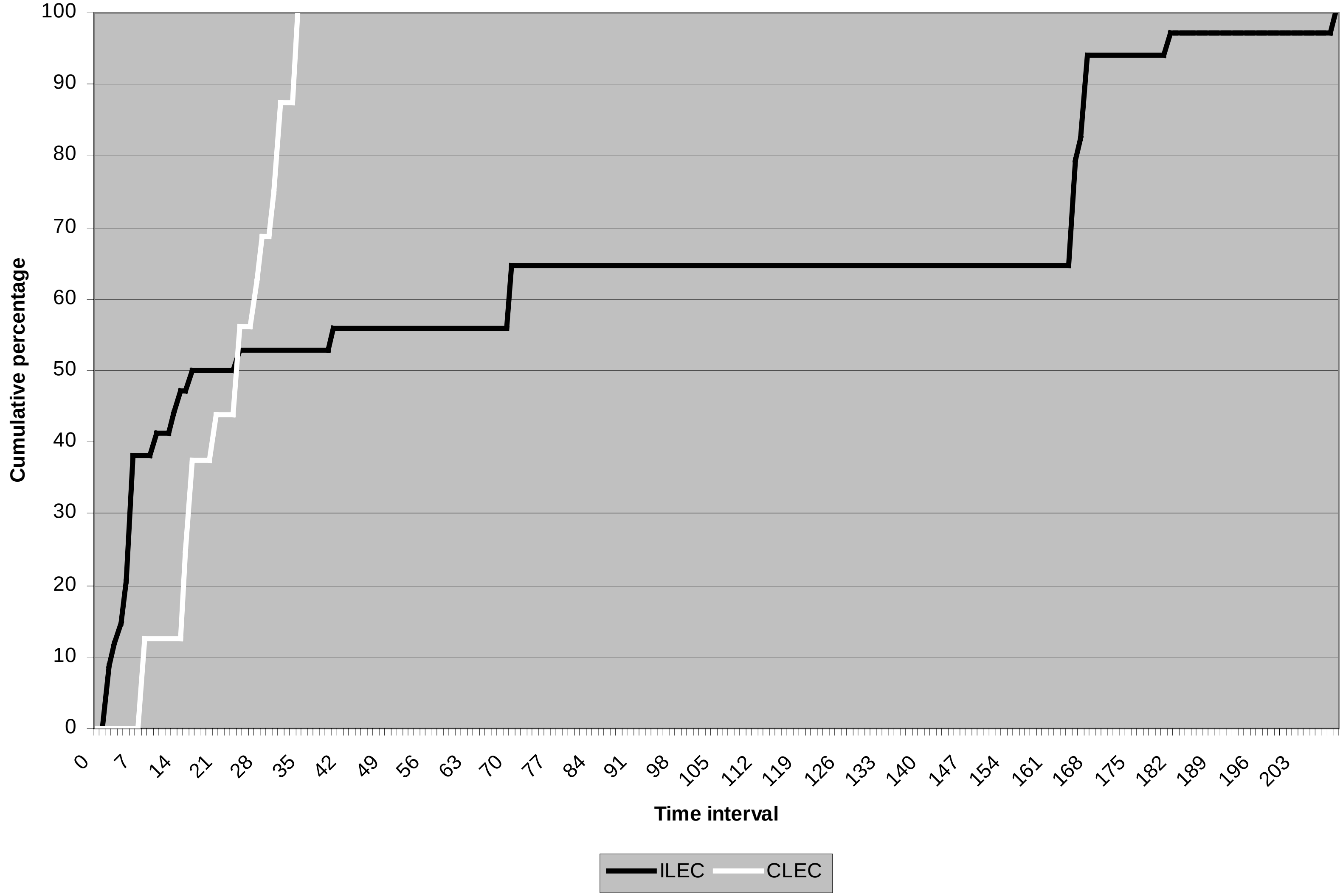
Cumulative distribution - Submeasure example 11



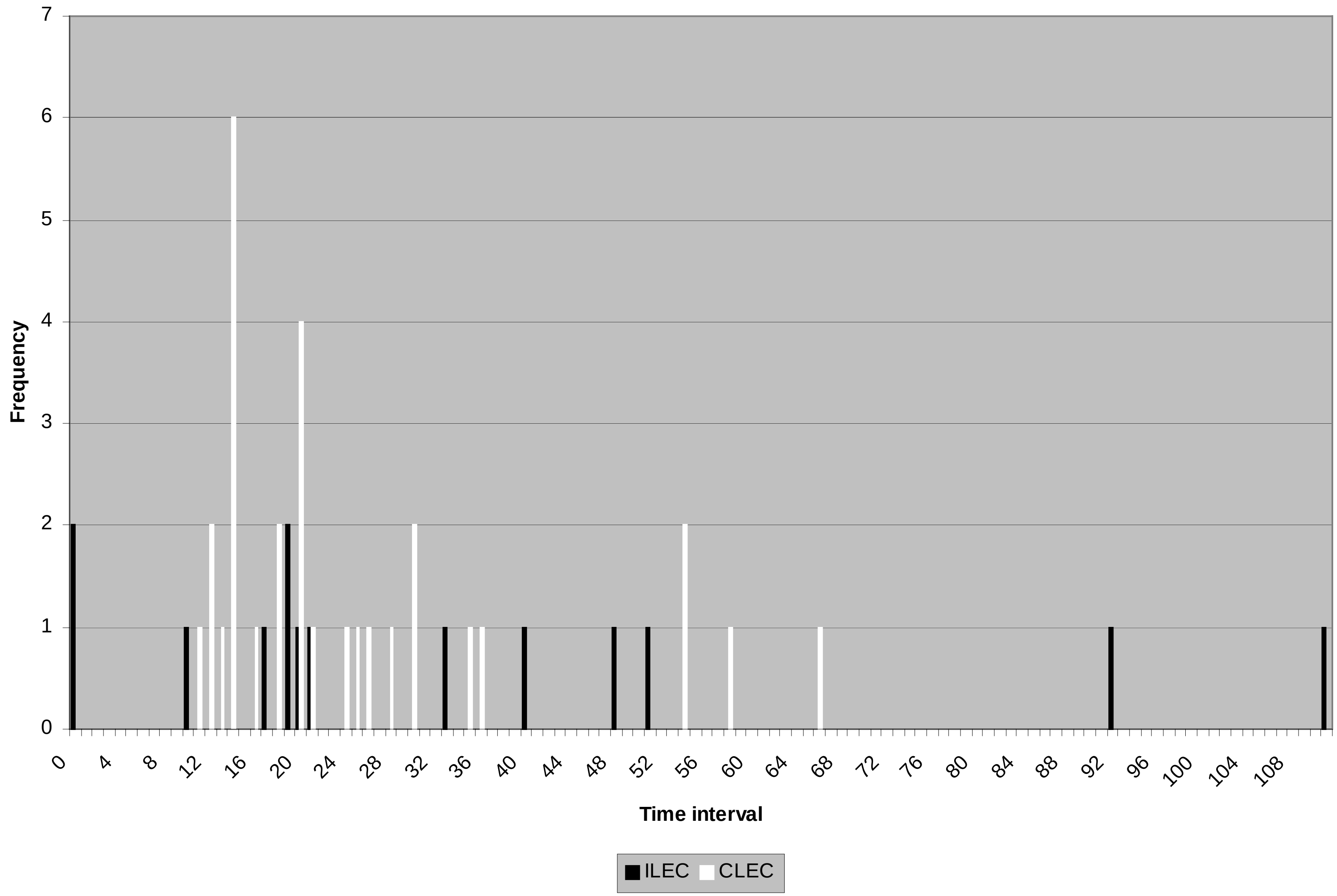
Frequency distribution - Submeasure example 12



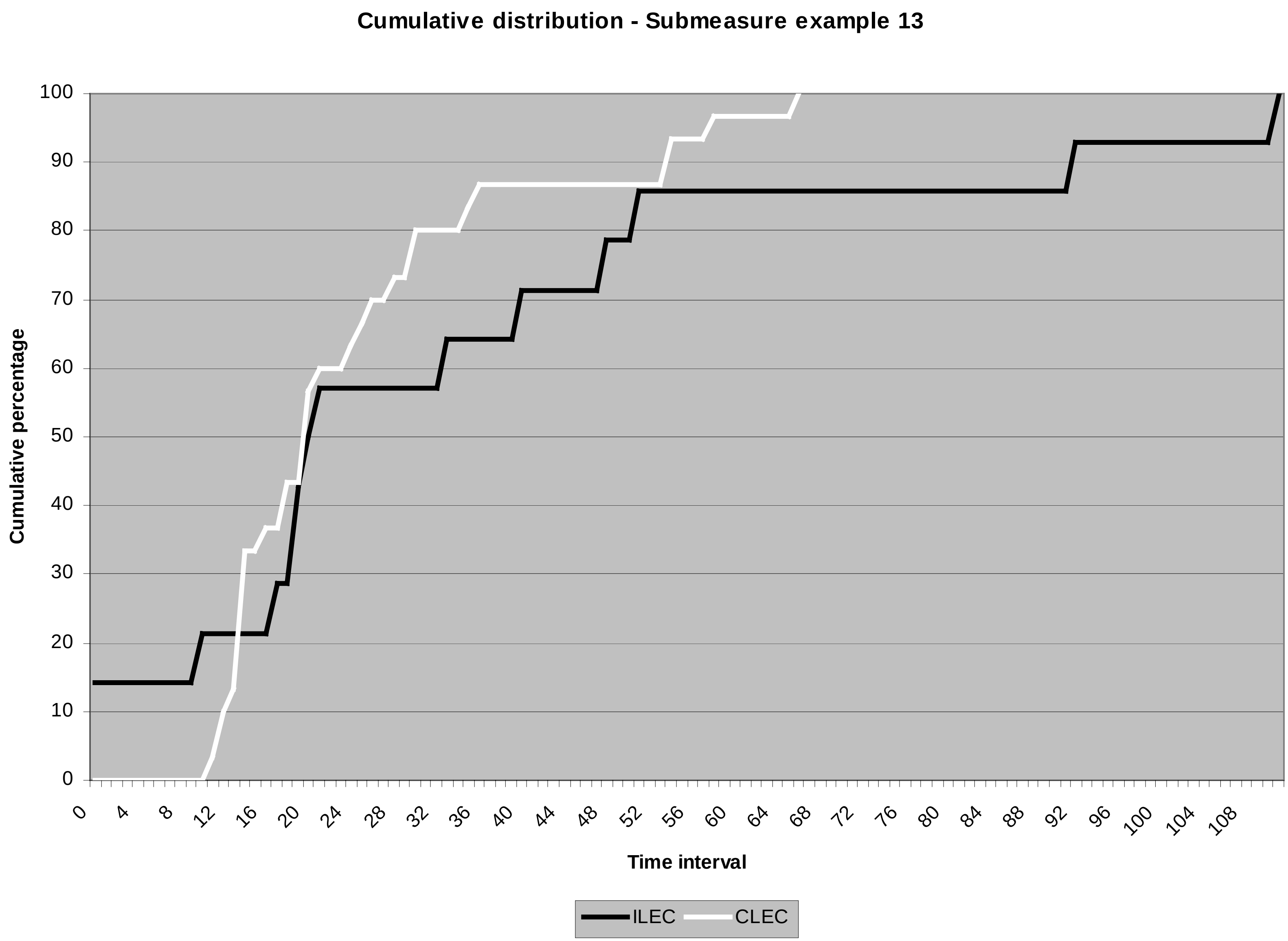
Cumulative distribution - Submeasure example 12

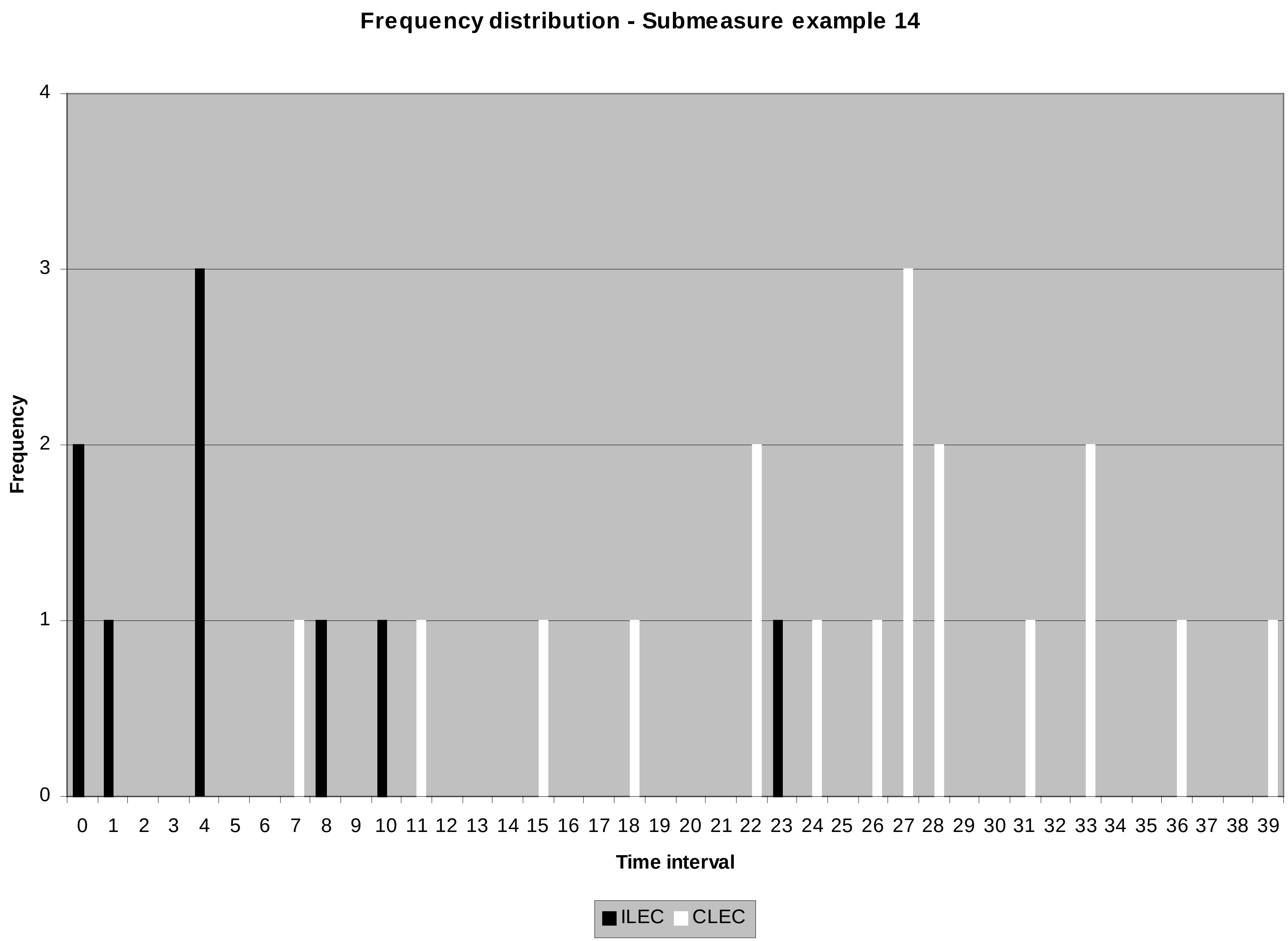


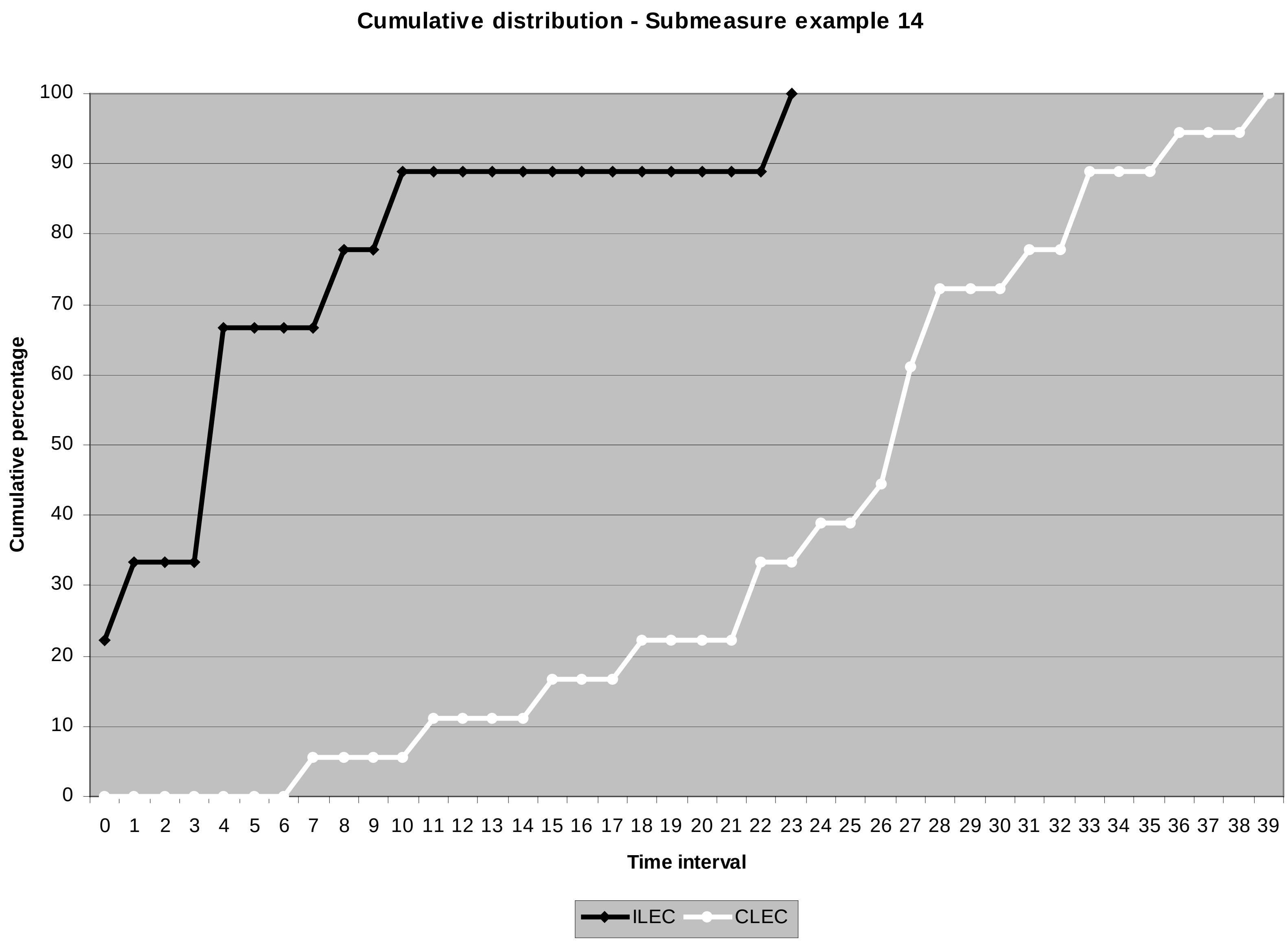
Frequency distribution - Submeasure example 13

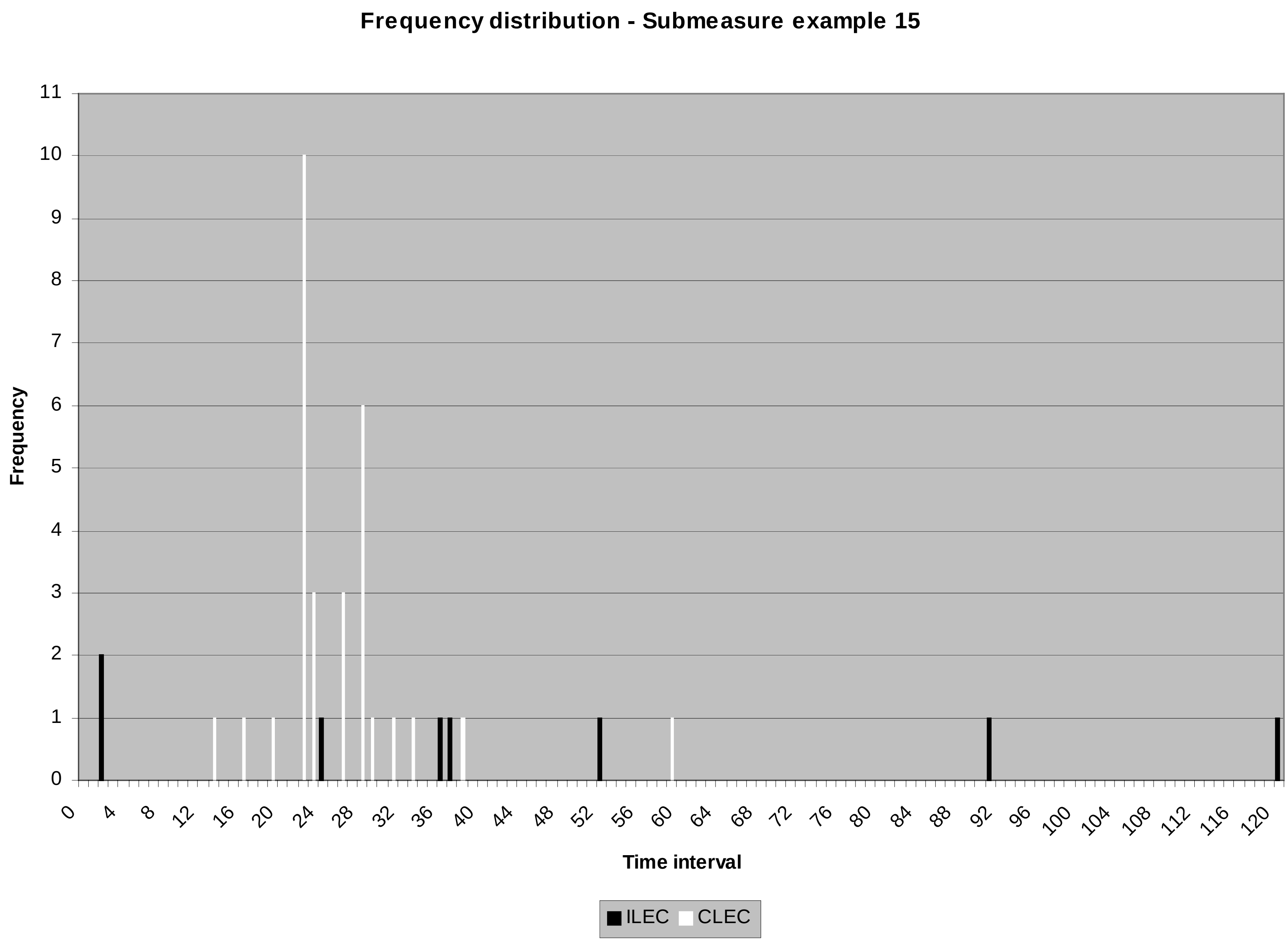


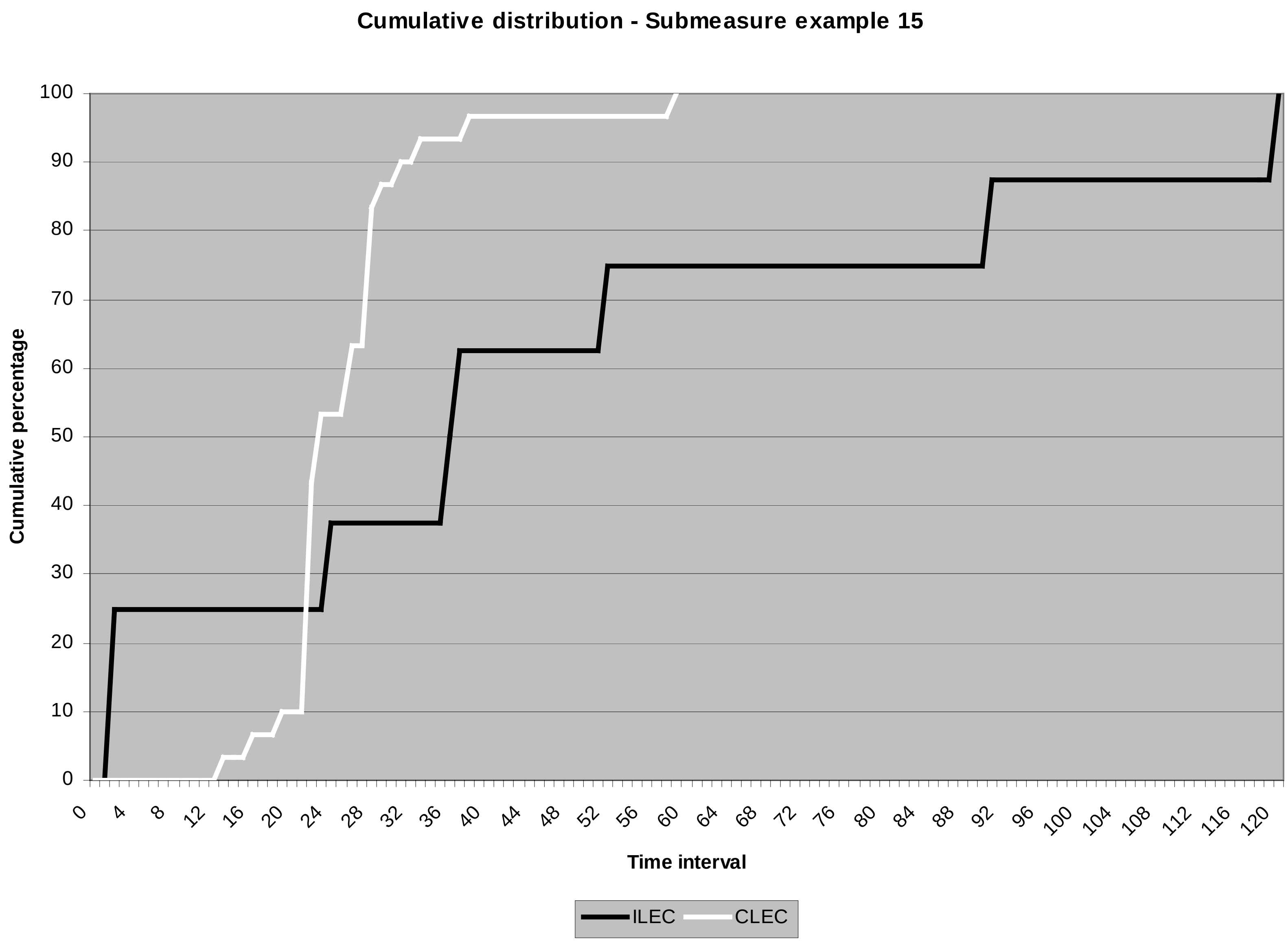
Cumulative distribution - Submeasure example 13



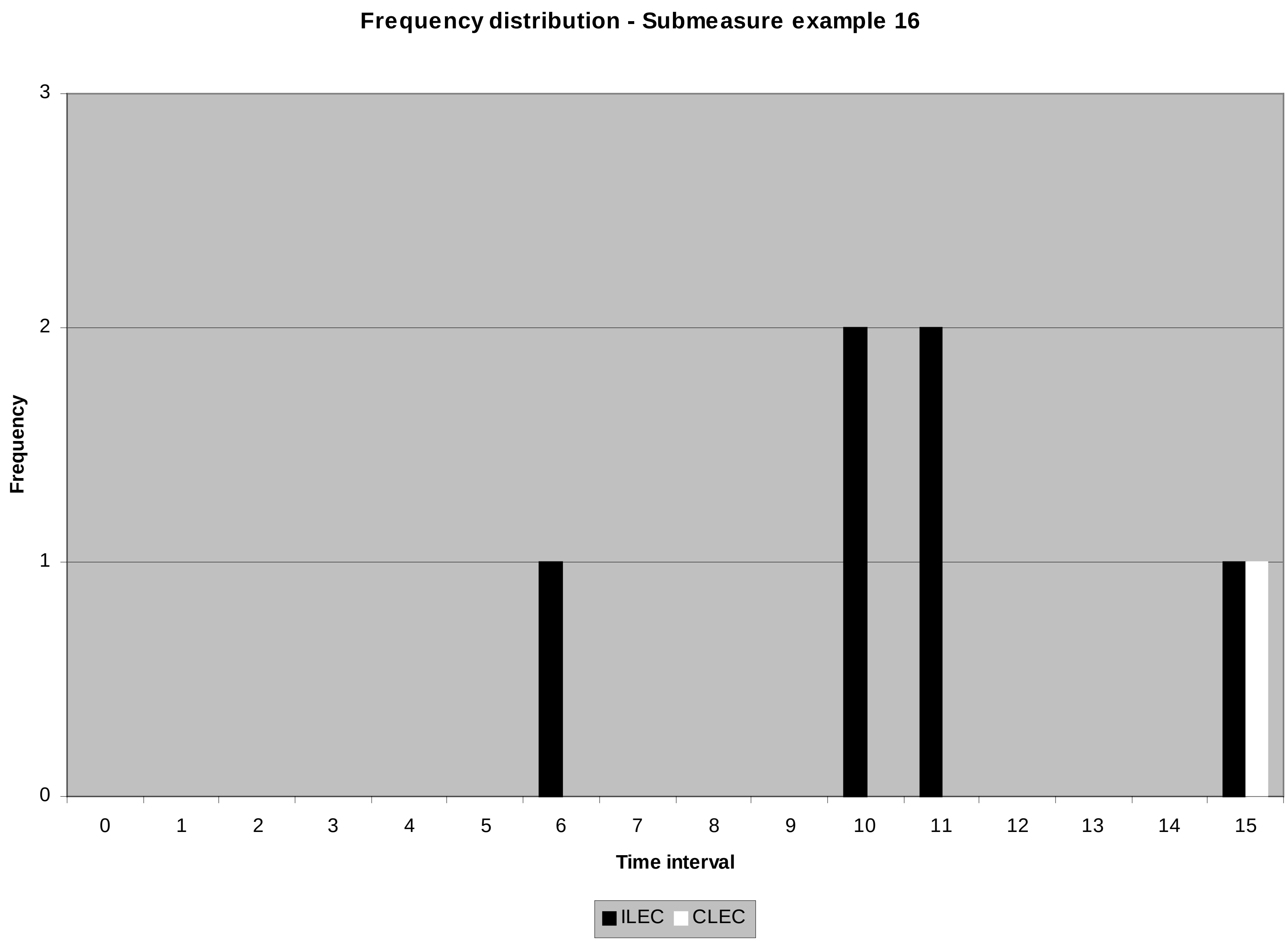

Frequency distribution - Submeasure example 14
Frequency
0
1
2
3
4
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
Time interval
ILEC
CLEC

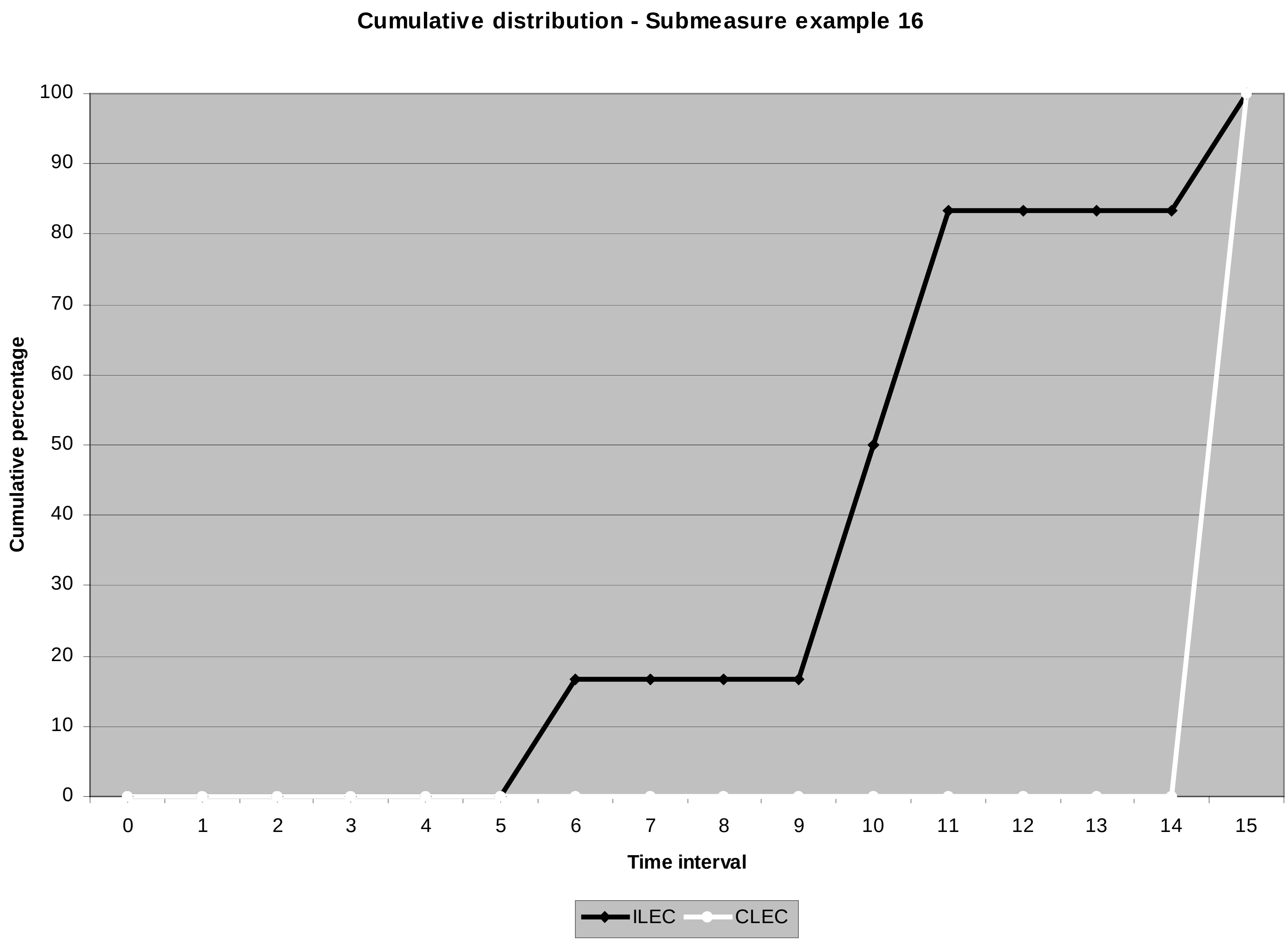

Cumulative distribution - Submeasure example 14
Cumulative percentage
100
90
80
70
60
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
Time interval
ILEC
CLEC


Frequency distribution - Submeasure example 15
Frequency
Time interval
ILEC
CLEC
0 1 2 3 4 5 6 7 8 9 10 11
0 4 8 12 16 20 24 28 32 36 40 44 48 52 56 60 64 68 72 76 80 84 88 92 96 100 104 108 112 116 120



Cumulative distribution - Submeasure example 15
Cumulative percentage
0
10
20
30
40
50
60
70
80
90
100
0
4
8
12
16
20
24
28
32
36
40
44
48
52
56
60
64
68
72
76
80
84
88
92
96
100
104
108
112
116
120
Time interval
ILEC
CLEC


Frequency distribution - Submeasure example 16
Frequency
3
2
1
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Time interval
ILEC
CLEC



Cumulative distribution - Submeasure example 16
Cumulative percentage
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Time interval
ILEC
CLEC

Mathcad worksheet: Investigation of the mean, standard deviation, skewness, and kurtosis of the sampling distribution of the mean before and after log transformations.

Set the parameters for the original process

$\mu := .8$ Process mean

$\sigma := 2.6$ Process standard deviation

Set sample sizes

$N := 100$ Sample size

Set additive constant for categorized distribution

$C := .4$ Additive constant

This simulation works by generating J samples of means for which the mean, standard deviation, skewness and kurtosis are calculated. This process is repeated K times and the means of the statistics on the sampling distributions are calculated.

Set number of samples for each simulation of the sampling distribution

$J := 100$ $j := 0..J-1$

Set number of simulations

$K := 100$ $k := 0..K-1$

The following calculate the log parameters of the distribution

$$m := \ln\left(\frac{\mu^2}{\sqrt{\mu^2+\sigma^2}}\right) \qquad m = -1.447$$

$$s := \sqrt{\ln\left[\frac{\left(\mu^2+\sigma^2\right)}{\mu^2}\right]} \qquad s = 1.565$$

The following function generates a log normal distribution using logl parameters.

$$g(n,m,s) := e^{\mathrm{rnorm}(n,m,s)}$$

This function generates means for four kinds of distributions:

1. lognormal
2. log of lognormal
3. categorized lognormal (into integers) with constant
4. log of categorized lognormal with constant

$$
f(j) := \left|
\begin{array}{l}
x \leftarrow g(N,m,s) \\
A_0 \leftarrow \mathrm{mean}(x) \\
A_1 \leftarrow \mathrm{mean}(\ln(x)) \\
y \leftarrow \mathrm{floor}(x) + C \\
A_2 \leftarrow \mathrm{mean}(y) \\
A_3 \leftarrow \mathrm{mean}(\ln(y)) \\
\text{return } A
\end{array}
\right.
$$

This function calculates statistics on each sample of sampling means and returns these statistics in a vector.

$$
\mathrm{sim}(k) := \left|
\begin{array}{l}
\text{for } \; j \in 0..J-1 \\
\quad X^{\langle j \rangle} \leftarrow f(j) \\
\text{for } \; h \in 0..3 \\
\quad \left|
\begin{array}{l}
Y_{4\cdot h+0} \leftarrow \mathrm{mean}\left[\left((X)^T\right)^{\langle h \rangle}\right] \\
Y_{4\cdot h+1} \leftarrow \mathrm{stdev}\left[\left((X)^T\right)^{\langle h \rangle}\right] \\
Y_{4\cdot h+2} \leftarrow \mathrm{skew}\left[\left((X)^T\right)^{\langle h \rangle}\right] \\
Y_{4\cdot h+3} \leftarrow \mathrm{kurt}\left[\left((X)^T\right)^{\langle h \rangle}\right]
\end{array}
\right. \\
\text{return } Y
\end{array}
\right.
$$

$$Y^{\langle k \rangle} := \mathrm{sim}(k)$$

Statistics for the distribution of sample means for the untransformed (original) distribution.

		Compare
Mean:	$\text{mean}\left[\left(Y^T\right)^{\langle 0\rangle}\right] = 0.794$	$\mu = 0.8$
Standard Deviation:	$\text{mean}\left[\left(Y^T\right)^{\langle 1\rangle}\right] = 0.247$	$\frac{\sigma}{\sqrt{N}} = 0.26$
Skewness:	$\text{mean}\left[\left(Y^T\right)^{\langle 2\rangle}\right] = 1.872$	
Kurtosis:	$\text{mean}\left[\left(Y^T\right)^{\langle 3\rangle}\right] = 7.846$	

Statistics for the distribution of sample means for the log transformed data.

		Compare
Mean:	$\text{mean}\left[\left(Y^T\right)^{\langle 4\rangle}\right] = -1.448$	$m = -1.447$
Standard Deviation:	$\text{mean}\left[\left(Y^T\right)^{\langle 5\rangle}\right] = 0.155$	$\frac{s}{\sqrt{N}} = 0.156$
Skewness:	$\text{mean}\left[\left(Y^T\right)^{\langle 6\rangle}\right] = -0.018$	
Kurtosis:	$\text{mean}\left[\left(Y^T\right)^{\langle 7\rangle}\right] = -0.011$	

Statistics for the distribution of sample means for the categorized data with an added constant of C.

		Compare
Mean:	$\text{mean}\left[\left(Y^T\right)^{\langle 8\rangle}\right] = 0.91$	$\mu = 0.8$
Standard Deviation:	$\text{mean}\left[\left(Y^T\right)^{\langle 9\rangle}\right] = 0.241$	$\frac{\sigma}{\sqrt{N}} = 0.26$
Skewness:	$\text{mean}\left[\left(Y^T\right)^{\langle 10\rangle}\right] = 1.963$	

Kurtosis: $\text{mean}\left[\left(Y^T\right)^{\langle 11\rangle}\right] = 8.344$

Statistics for the distribution of sample means for the logs of the categorized data with an added constant of C.

Compare

Mean: $\text{mean}\left[\left(Y^T\right)^{\langle 12\rangle}\right] = -0.603$ $m = -1.447$

Standard Deviation: $\text{mean}\left[\left(Y^T\right)^{\langle 13\rangle}\right] = 0.071$ $\frac{s}{\sqrt{N}} = 0.156$

Skewness: $\text{mean}\left[\left(Y^T\right)^{\langle 14\rangle}\right] = 0.206$

Kurtosis: $\text{mean}\left[\left(Y^T\right)^{\langle 15\rangle}\right] = 1.557 \cdot 10^{-3}$

Skewness and kurtosis variability of theoretical sampling means								
Raw data parameters					Sampling mean distribution			
					Untransformed data		Transformed data	
Sample size	SD	M	Cases	Constant	Skewness	Kurtosis	Skewness	Kurtosis
1000	2.6	0.5	24	0.3	1.7	7.1	0.1	0.2
					1.7	6.7	0.1	0.1
					1.6	5.9	0.1	-0.1
					1.6	5.9	0.1	0.0
					1.8	7.7	0.2	0.1
1000	2.5	0.8	129	0.4	0.9	2.2	0.1	-0.1
					0.7	1.1	0.2	-0.2
					0.9	1.6	0.1	0.0
					0.7	1.0	0.0	-0.1
					1.2	4.3	0.1	0.3
100	2.5	0.8	129	0.4	1.8	7.0	0.2	0.1
					1.7	5.3	0.2	0.1
					1.8	6.1	0.3	0.0
					2.1	9.0	0.2	0.1
					1.6	5.0	0.2	0.0
1000	6	2.9	70	0.5	0.3	0.4	-0.1	-0.2
					0.5	0.8	0.0	-0.1
					0.4	0.5	-0.1	0.0
					0.9	3.5	0.1	-0.1
					0.5	0.5	0.0	-0.1
100	6	2.9	70	0.5	1.4	4.9	0.1	0.0
					0.9	1.4	0.2	-0.2
					3.1	34.3	0.1	0.0
					1.7	8.9	0.0	0.1
					2.6	24.1	0.0	0.5
100	7	6	131	0.5	0.6	0.7	0.0	0.2
					0.6	1.1	0.1	-0.1
					0.4	0.3	0.0	0.1
					0.3	0.0	-0.1	0.0
					0.6	1.1	-0.1	0.1
30	7	6	131	0.5	1.2	3.3	0.0	0.0
					0.9	1.5	0.0	-0.1
					1.0	2.2	-0.1	-0.2
					1.1	2.3	0.0	-0.2
					0.8	1.1	-0.1	0.2
1000	16	3.1	24	0.4	1.0	1.9	0.0	-0.2
					3.2	32.2	0.0	-0.1
					1.5	4.3	0.1	0.4
					1.3	3.0	0.0	0.0
					1.6	7.1	-0.1	0.1

100	25	13	29	0.5	0.9	1.5	0.0	-0.2
					2.0	11.4	0.0	0.1
					0.9	1.6	0.1	0.6
					1.1	2.3	0.1	0.3
					1.1	3.0	-0.1	0.0

Mathcad worksheet: Investigation of the added constant used with the log transformation on a categorized lognormal distribution

Set the parameters for the original process

$\mu := .9$ Process mean

$\sigma := 4.7$ Process standard deviation

Set size of sample for investigating distribution

$N := 100000$

The following calculate the log parameters of the distribution

$$m := \ln\left(\frac{\mu^2}{\sqrt{\mu^2 + \sigma^2}}\right) \qquad m = -1.776$$

$$s := \ln\left[\frac{\left(\mu^2 + \sigma^2\right)}{\mu^2}\right]^{\left(\frac{1}{2}\right)} \qquad s = 1.828$$

The following function generates a log normal distribution using log parameters.

$g(n,m,s) := e^{\mathrm{rnorm}(n,m,s)}$

$x := g(N,m,s)$

Theoretical mean $\mu = 0.9$

Emprical mean $\mathrm{mean}(x) = 0.886$

Theoretical standard deviation $\sigma = 4.7$

Emprical standard deviation $\mathrm{stdev}(x) = 4.157$

$l := \mathrm{ceil}(\max(x)) + 1 \qquad l = 622$

$i := 0 .. I$

$int_i := i$

$y := \text{hist}(int, x)$ $\qquad \sum y = 1 \cdot 10^5$

$ii := 0 .. I - 1$

$w_{ii} := ii$

The following solves for a, that value which, when added to the bottom end of each interval, best recreates the emprical mean

$K := \frac{y \cdot w}{N}$ $\qquad K = 0.637$

$C := (K - \text{mean}(x))^2$ $\qquad B := 2 \cdot (K - \text{mean}(x))$

$\alpha := \frac{-B + \sqrt{B^2 - 4 \cdot C}}{2}$ $\qquad \alpha = 0.249$

$z_{ii} := w_{ii} + \alpha$

$\text{length}(y) = 622$

$\text{length}(z) = 622$

Mean using added constant $\qquad \frac{y \cdot z}{N} = 0.886$

Standard deviation using added constant $\qquad \sqrt{\frac{y \cdot z^2}{N - 1} - \left(\frac{y \cdot z}{N}\right)^2} = 4.115$

$\text{mean}(\ln(x)) = -1.775$ $\qquad$ (Compare to the log parameters above)

$\frac{y \cdot \ln(z)}{N} = -1.012$

Skewness and kurtosis of performance result distributions				
	Raw score	Log(x+0.5)	Log(x+0.4)	Log(x+0.3)
Ex. 1				
N	179254	179254	179254	179254
Mean	1.18	-0.1128	-0.2808	-0.4948
Median	0	-0.6931	-0.9163	-1.204
Skewness	21.002	1.331	1.291	1.244
Kurtosis	1430.297	0.28	0.126	-0.046
Ex. 2				
N	23608	23608	23608	23608
Mean	1.6	0.1008	-3.81E-02	-0.2122
Median	0	-0.6931	-0.9163	-1.204
Skewness	15.411	1.053	0.942	0.817
Kurtosis	503.008	0.71	0.3	-0.131
Ex. 3				
N	19943	19943	19943	19943
Mean	6.91	1.8604	1.8405	1.8193
Median	6	1.8718	1.8563	1.8405
Skewness	12.106	-1.27	-1.513	-1.818
Kurtosis	271.231	9.202	9.945	11.092
Ex. 4				
N	17951	17951	17951	17951
Mean	0.9215	-0.1906	-0.3589	-0.5723
Median	0	-0.6931	-0.9163	-1.204
Skewness	28.256	1.634	1.528	1.412
Kurtosis	1745.695	2.331	1.773	1.188
Ex. 5				
N	17940	17940	17940	17940
Mean	2.76	0.8407	0.783	0.7187
Median	2	0.9163	0.8755	0.8329
Skewness	15.569	0.666	0.499	0.278
Kurtosis	590.337	1.57	1.473	1.453
Ex. 6				
N	11864	11864	11864	11864
Mean	1.3988	5.55E-02	-9.15E-02	-0.277
Median	0	-0.6931	-0.9163	-1.204
Skewness	9.105	0.925	0.861	0.789
Kurtosis	184.585	-0.336	-0.538	-0.755

Skewness and kurtosis of performance result distributions				
Ex. 7				
N	9149	9149	9149	9149
Mean	1.2922	-8.06E-02	-0.2426	-0.4484
Median	0	-0.6931	-0.9163	-1.204
Skewness	19.716	1.362	1.289	1.208
Kurtosis	661.21	0.903	0.585	0.246
Ex. 8				
N	6827	6827	6827	6827
Mean	2.4837	0.7652	0.7077	0.6453
Median	1	0.4055	0.3365	0.2624
Skewness	10.337	1.493	1.364	1.187
Kurtosis	198.575	3.221	3.021	2.868
Ex. 9				
N	6340	6340	6340	6340
Mean	3.05	0.8676	0.8113	0.7491
Median	1	0.4055	0.3365	0.2624
Skewness	5.295	0.993	0.855	0.668
Kurtosis	48.225	1.505	1.414	1.377
Ex. 10				
N	771	771	771	771
Mean	8.18	1.7666	1.7302	1.6875
Median	7	2.0149	2.0015	1.9879
Skewness	6.917	-0.998	-1.094	-1.214
Kurtosis	105.315	0.542	0.749	1.03
Ex. 11				
N	538	538	538	538
Mean	7.89	1.7286	1.6883	1.6402
Median	7	2.0149	2.0015	1.9879
Skewness	1.81	-1.096	-1.177	-1.277
Kurtosis	8.242	0.408	0.579	0.802
Ex. 12				
N	34	34	34	34
Mean	71.6176	3.2922	3.2818	3.2712
Median	20	3.001	2.9959	2.9908
Skewness	0.525	-0.017	-0.023	-0.03
Kurtosis	-1.623	-1.712	-1.705	-1.698

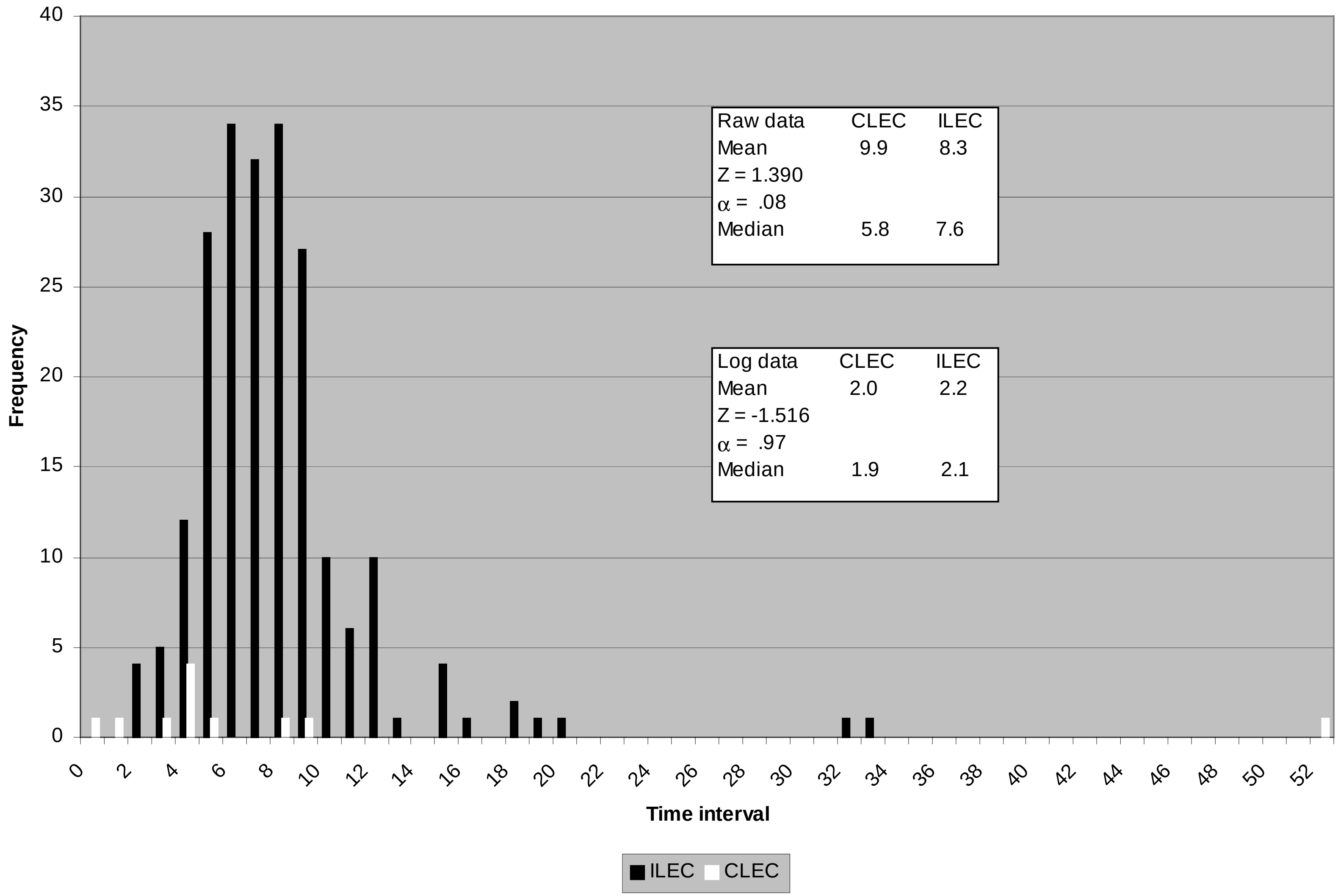
Skewness and kurtosis of performance result distributions				
Ex. 13				
N	14	14	14	14
Mean	34.3571	2.8667	2.8315	2.7871
Median	20.5	3.0442	3.0395	3.0347
Skewness	1.389	-1.54	-1.597	-1.664
Kurtosis	1.484	1.906	2.033	2.184
Ex. 14				
N	9	9	9	9
Mean	6	1.2422	1.1746	1.0919
Median	4	1.5041	1.4816	1.4586
Skewness	1.874	-0.407	-0.492	-0.594
Kurtosis	3.97	-0.763	-0.73	-0.671
Ex. 15				
N	8	8	8	8
Mean	47.5	3.244	3.2322	3.2201
Median	40.5	3.7135	3.7111	3.7086
Skewness	0.732	-1.041	-1.048	-1.056
Kurtosis	-0.114	-0.426	-0.419	-0.412
Ex. 16				
N	6	6	6	6
Mean	10.5	2.3667	2.3569	2.347
Median	10.5	2.3969	2.3877	2.3785
Skewness	0	-0.912	-0.92	-0.929
Kurtosis	2.086	2.601	2.611	2.62

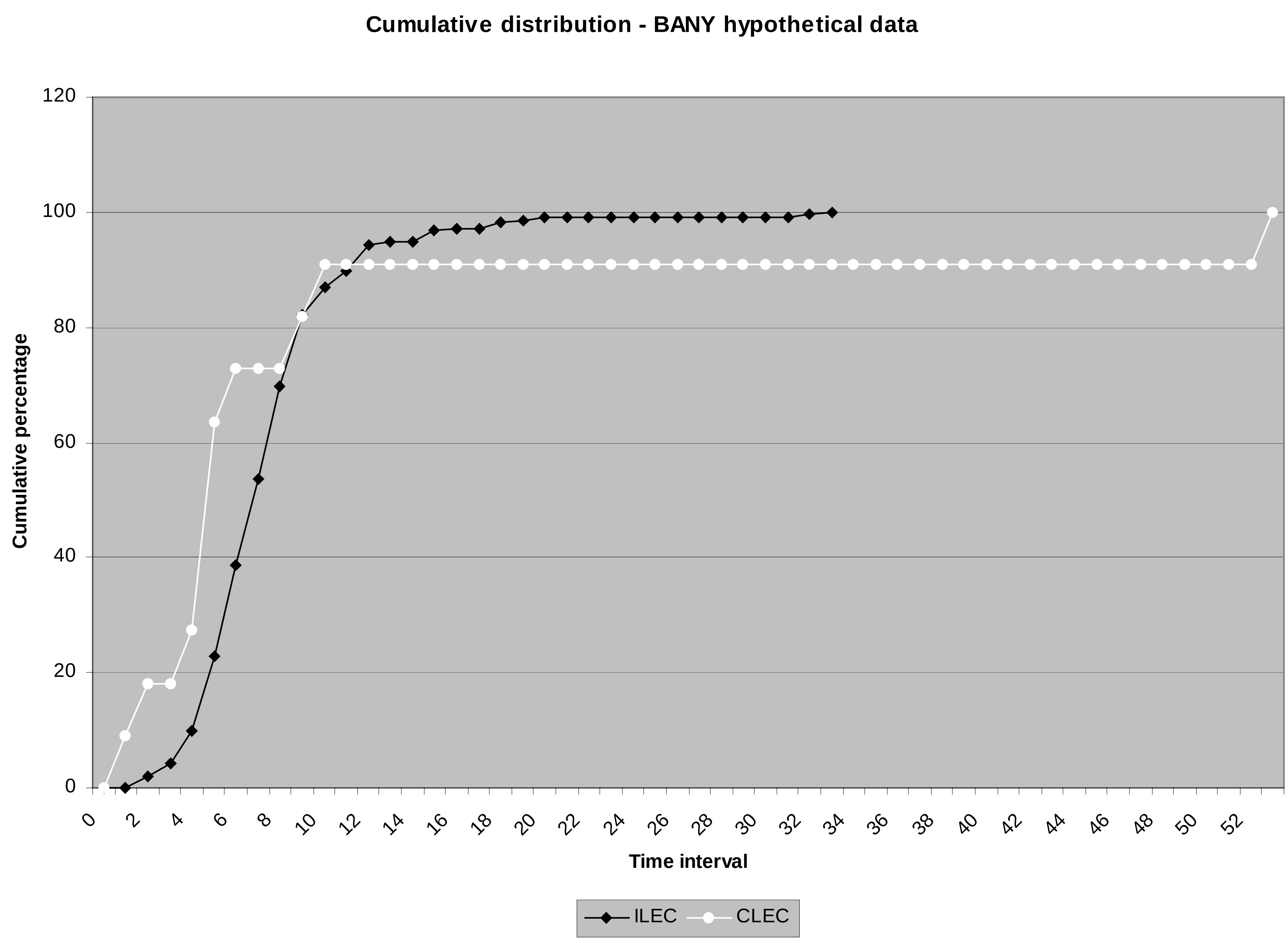
Sensitivity analysis: Effects of transformations on alpha estimates

Ex. No.	Constant used	N_i	M_i	SD_i	N_c	M_c	Z	α	Theoretical constant
1	no trnsfmn	179254	1.18	2.97	4296	0.149	-22.485	1.00	
1	0.5	179254	-0.1128	0.9696	4296	-0.5295	-27.837	1.00	
1	0.4	179254	-0.2808	1.0537	4296	-0.7297	-27.594	1.00	0.44
1	0.3	179254	-0.4948	1.653	4296	-0.9855	-19.228	1.00	
2	no trnsfmn	23608	1.6	4.8	21	3.52	1.832	0.0335	
2	0.5	23608	0.1008	0.9636	21	0.7646	3.155	0.0008	0.48
2	0.4	23608	-0.0381	1.0474	21	0.6724	3.107	0.0009	
2	0.3	23608	-0.2122	1.1602	21	0.558	3.041	0.0012	
4	no trnsfmn	17951	0.9215	3.4631	276	0.337	-2.783	0.997	
4	0.5	17951	-0.1906	0.8292	276	-0.3251	-2.674	0.996	
4	0.4	17951	-0.3589	0.907	276	-0.4968	-2.507	0.994	0.36
4	0.3	17951	-0.5723	1.012	276	-0.7131	-2.294	0.989	
5	no trnsfmn	17940	2.76	4.69	302	1.755	-3.693	0.9999	
5	0.5	17940	0.8407	0.739	302	0.6969	-3.353	0.9996	0.50
5	0.4	17940	0.783	0.7783	302	0.6348	-3.282	0.9995	
5	0.3	17940	0.7147	0.8282	302	0.5651	-3.113	0.9991	
7	no trnsfmn	9149	1.2922	4.3519	30	5.4	5.162	0.00000013	
7	0.5	9149	-0.08056	0.9588	30	1.3947	8.414	0.00000000	0.40
7	0.4	9149	-0.2426	1.0421	30	1.3477	8.345	0.00000000	
7	0.3	9149	-0.4484	1.1535	30	1.2916	8.249	0.00000000	
9	no trnsfmn	6340	3.05	4.9	714	3.1387	0.459	0.32	
9	0.5	6340	0.8676	0.782	714	1.1988	10.729	0.00	0.50
9	0.4	6340	0.8113	0.8171	714	1.1639	10.932	0.00	
9	0.3	6340	0.7491	0.8608	714	1.1269	11.118	0.00	
10	no trnsfmn	771	8.18	7.95	179	8.45	0.409	0.341	
10	0.5	771	1.7666	1.0351	179	1.96	2.252	0.012	0.50
10	0.4	771	1.7302	1.0883	179	1.9397	2.320	0.010	
10	0.3	771	1.6875	1.1571	179	1.9181	2.402	0.008	
11	no trnsfmn	538	7.89	6.33	115	9.0696	1.814	0.0351	
11	0.5	538	1.7286	1.0742	115	2.0628	3.028	0.0013	0.50
11	0.4	538	1.6883	1.1347	115	2.0468	3.075	0.0011	
11	0.3	538	1.6402	1.2133	115	2.0303	3.130	0.0009	

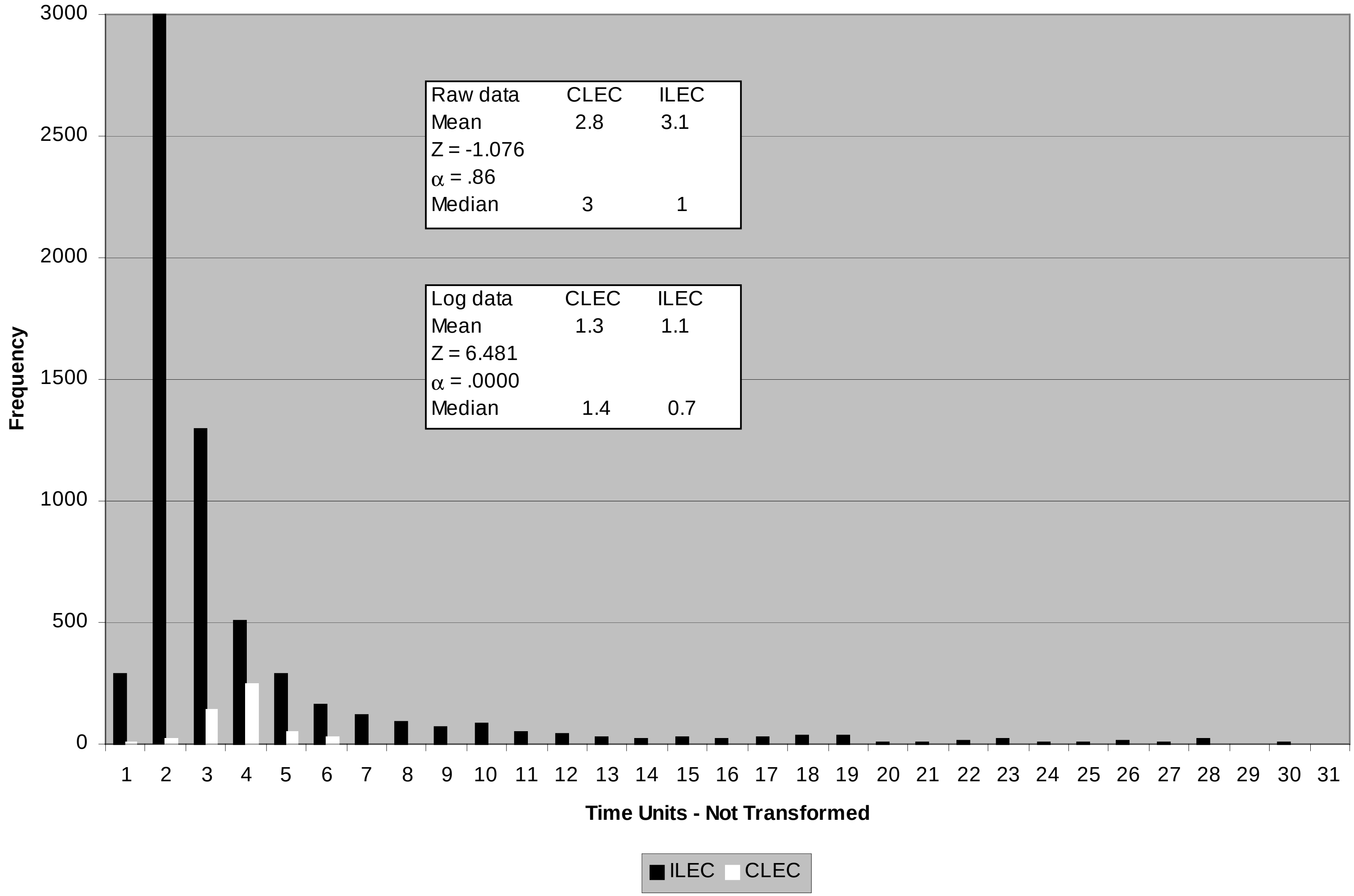
Ex. No.	Constant used	N_i	M_i	SD_i	N_c	M_c	Z	α	Theoretical constant
13	no trnsfmn	14	34.3571	32.5874	30	25.03	-0.884	0.80	
13	0.5	14	2.8667	1.6447	30	3.1086	0.454	0.33	0.50
13	0.4	14	2.8315	1.7188	30	3.1037	0.489	0.32	
13	0.3	14	2.7871	1.8152	30	3.0987	0.530	0.30	
14	no trnsfmn	9	6	7.2284	18	25.2222	6.514	0.00009	
14	0.5	9	1.2422	1.3276	18	3.1782	3.572	0.00364	0.50
14	0.4	9	1.1746	1.4104	18	3.1736	3.472	0.00421	
14	0.3	9	1.0919	1.5187	18	3.1689	3.350	0.00504	
15	no trnsfmn	8	47.5	41.127	30	25.77	-1.328	0.887	
15	0.5	8	3.244	1.5084	30	3.2332	-0.018	0.507	0.50
15	0.4	8	3.2322	1.5258	30	3.2291	-0.005	0.502	
15	0.3	8	3.2201	1.5439	30	3.225	0.008	0.497	

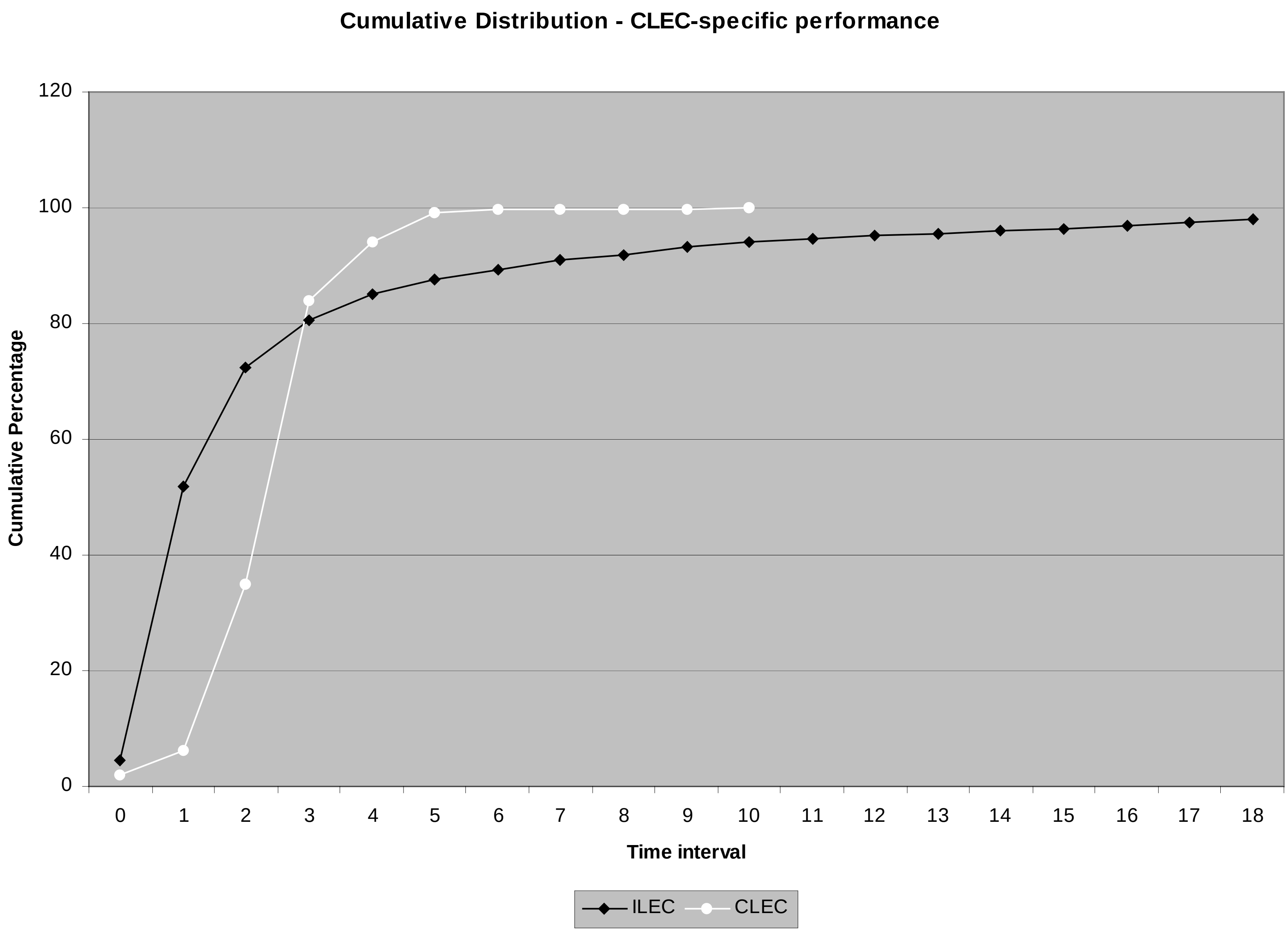
Distributions - BANY hypothetical data




Cumulative distribution - BANY hypothetical data
120
100
80
60
40
20
0
Cumulative percentage
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52
Time interval
ILEC
CLEC

Frequency distribution - CLEC-specific performance




Cumulative Distribution - CLEC-specific performance
120
100
80
60
40
20
0
Cumulative Percentage
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Time interval
ILEC
CLEC

Appendix K

Benchmark Small Sample Adjustment Tables

Benchmark Small Sample Adjustment Tables: Derivation and Application Methods

Introduction

This appendix describes the method and rationale for the construction and application of benchmark small-sample adjustment tables. These tables provide allowances for performance results that fail benchmarks because of problems posed by small samples.

For performance to pass a strictly applied benchmark, results for small samples can often require perfect performance even though the benchmark allows less than perfect performance, such as 90 percent of OSS tasks completed on time. Additionally, with infrequent exception, strict application requires performance actually higher than the benchmark. This poses somewhat of a dilemma because, similar to the "alpha vs. beta" problem in statistical testing, solutions for one situation cause problems for the other situation. The parties have somewhat irreconcilable interests. The ILECs are concerned with erroneous decisions under conditions when an operational process actually allows "a meaningful opportunity to compete" as defined by the benchmark. The CLECs are concerned with erroneous decisions under conditions when an operational process actually does not allow "a meaningful opportunity to compete" as defined by the benchmark. The following examples illustrate two types of problems.

Example 1: Problems under "meaningful opportunity" (passing) conditions.

These problems exist when aggregate performance for all CLECs *allows* a meaningful opportunity to compete as illustrated by the following results for a 90 percent benchmark. Twenty CLECs submit five orders each for a total of 100 orders. We assume 10 "delinquent orders," that is, orders with necessary OSS tasks not completed within the benchmark time interval criterion. The process is in parity because 90 percent of the orders are completed within the time criterion. However, between 2 and 10 CLECs will *fail* the benchmark. The performance for at least two CLECs will fail as illustrated by the following outcome:

a. Two CLECs have no orders completed within the benchmark timeframe (zero percent performance result), accounting for the 10 delinquent orders (2 CLECs x 5 orders each).

The performance for as many as ten CLECs could fail the benchmark as illustrated by the following outcome:

b. Ten CLECs each have one order not completed within the benchmark timeframe (80 percent performance result), accounting for the 10 delinquent orders (10 CLECs x 1 order each).

These situations illustrate the logical underpinning of the small sample adjustment tables. In this illustration, a small sample table would allow one additional miss per CLEC result. This would allow the more likely outcomes (one delinquent order per CLEC) to be deemed in parity consistent with the 90 percent aggregate process performance level.

Example 2: Problems under “no meaningful opportunity” (failure) conditions:

These problems exist when aggregate performance for all CLECs *does not allow* a meaningful opportunity to compete, as illustrated with the following results, again for a 90 percent benchmark. Twenty CLECs submit five orders each for a total of 100 orders. We assume 20 delinquent orders. The process is failing because only 80 percent of orders are within the time criterion. However, between 10 and 20 (all) CLECs will *pass* the benchmark when a small sample table is used allowing one delinquent order per CLEC. The performance for at least ten CLECs will pass, as illustrated by the following outcome:

a. Ten CLECs have two orders not completed within the benchmark timeframe (sixty percent performance result), accounting for the 20 delinquent orders (10 CLECs x 2 delinquent orders each). The remaining ten CLECs would have no delinquent orders and thus would pass the benchmark.

The performance for as many as all twenty CLECs could pass the benchmark as illustrated by the following outcome:

b. Twenty CLECs each have one order not completed within the benchmark timeframe (80 percent performance result), accounting for the 20 delinquent orders (20 CLECs x 1 order each). However, all CLECs pass the benchmark because the small sample adjustment tables allow one "miss" each.

When strictly applied, the benchmarks actually require performance higher than the nominally specified percentage level as demonstrated in Table 1.

Table 1

(1) Sample size	(2) "Misses" permitted: Absolute benchmark	(3) Effective percentage: Absolute benchmark	(4) "Misses" permitted: Adjusted* benchmark	(5) Effective percentage: Adjusted* benchmark	(6) Sample size percentage	(7) Weighted effective percentage: Absolute benchmark	(8) Weighted effective percentage: Adjusted* benchmark
1	0	100.0%	1	0.0%	15.9	15.9	0.0
2	0	100.0%	1	50.0%	8.7	8.7	4.4
3	0	100.0%	1	66.7%	5.9	5.9	3.9
4	0	100.0%	1	75.0%	8.5	8.5	6.3
5	0	100.0%	1	80.0%	3.1	3.1	2.5
6	0	100.0%	1	83.3%	5.9	5.9	4.9
7	0	100.0%	1	85.7%	4.4	4.4	3.7
8	0	100.0%	1	87.5%	3.8	3.8	3.4
9	0	100.0%	1	88.9%	3.6	3.6	3.2
10	1	90.0%	1	90.0%	2.6	2.3	2.1
11	1	90.9%	2	81.8%	2.8	2.6	2.1
12	1	91.7%	2	83.3%	2.3	2.1	1.8
13	1	92.3%	2	84.6%	2.1	1.9	1.6
14	1	92.9%	2	85.7%	3.6	3.3	2.9
15	1	93.3%	2	86.7%	2.8	2.6	2.3
16	1	93.8%	2	87.5%	2.3	2.2	1.9
17	1	94.1%	2	88.2%	2.8	2.7	2.3
18	1	94.4%	2	88.9%	3.1	2.9	2.6
19	1	94.7%	2	89.5%	1.8	1.7	1.5
20	2	90.0%	2	90.0%	2.3	2.1	1.9
21	2	90.5%	3	85.7%	1.5	1.4	1.2
22	2	90.9%	3	86.4%	1.8	1.6	1.4
23	2	91.3%	3	87.0%	1.0	0.9	0.8
24	2	91.7%	3	87.5%	1.8	1.6	1.4
25	2	92.0%	3	88.0%	1.8	1.7	1.5
26	2	92.3%	3	88.5%	0.5	0.5	0.4
27	2	92.6%	3	88.9%	0.3	0.2	0.2
28	2	92.9%	3	89.3%	0.8	0.7	0.6
29	2	93.1%	3	89.7%	1.3	1.2	1.1
30	3	90.0%	3	90.0%	1.0	0.9	0.8
	Average =	94.5%		81.8%	Sum = 100.0	96.9	64.7

* Adjusted by adding one (1) additional permitted miss to each sample size result unless the original effective percentage equals the 90% nominal percentage.

Table 1 shows the effective percentage for the 90 percent benchmark for sample sizes of one (1) to thirty (30) both for an absolute benchmark application and an application allowing one additional “miss.” For example, for a sample size of 19, since two “misses” out of 19 orders equals 89.5 percent on-time performance, it fails the benchmark. Therefore with an absolute application only one miss is allowed (column 2). Thus a performance result of at least 18 on-time orders out of 19 is required to pass – effectively a 94.7 percent performance requirement (column 3). On average, for sample sizes of one to thirty, a 90 percent benchmark requires 94.5 percent on-time performance to pass the benchmark. On the other hand, if one additional miss (column 4) was allowed so that the effective percentage was never greater than the nominal benchmark percentage, then this “adjusted” benchmark would only require 81.8 percent on-time performance for this range of samples (column 5). When these averages are adjusted for the fact that some sample sizes are more numerous than others, the absolute benchmark has an effective percentage of 96.9 percent, whereas the adjusted benchmark has an effective percentage of 64.7 percent.[1]

The above discussion addresses what the parties have described as the “granularity” problem with small sample benchmark application. That is, when failures are in integer increments, and the integer is larger than the permissible percentage of misses, then performance higher than the benchmark is required. The best illustration of this phenomenon is Example 1b above, where it is not possible to avoid identifying failures even though aggregate performance passes the benchmark. However, since benchmark adjustment can only be accomplished in integer increments, then adjustments can result in failing performance being identified as passing. This phenomenon is best illustrated in example 2b above, where no failures are identified even though aggregate performance fails the benchmark.

The granularity problem is distinguishable from a “random variation” problem. A random variation problem is illustrated in the following example:

---

[1] For example, there is approximately twice the number of results with a sample size of one as there is with a sample size of two (15.9 versus 8.7 percent, respectively; see column 6). Consequently, the average effective percentage across all sample sizes will be affected more by the effective percentage for sample sizes of one than by sample sizes of two. Table 1 columns 7 and 8 effectively account for this relative difference. This table is for illustration purposes only.

Example 3:

With ten orders each, altogether ten CLECs have 100 orders. Overall there are ten delinquent orders for a passing percentage of 90 percent. If each CLEC has one delinquent order, their performance results will all pass the benchmark. However, because of the "luck of the draw" it is unlikely that the delinquent orders will be distributed equally across all CLECs. Instead, two CLECs have two delinquent orders each, six CLECs have one delinquent order each, and two CLECs have no delinquent orders.

Thus, the performance results for two CLECs fail even though the overall performance passes the benchmark. Pacific has proposed statistical testing of benchmarks to mitigate the effects of random variation in this situation, which is analogous to the "Type I error" situation for parity measure assessment.

However, the parties have agreed to address the granularity problem, but not the random variation problem. The only exception is that the parties have agreed to use a statistical method to create the small sample adjustment tables. This appendix proceeds within these guidelines. A complete statistical treatment is not established. The granularity problem is addressed through statistically developed small sample adjustment tables.

Proposals

In summary, while small sample table allowances alleviate one problem, they cause another problem. Analyses of ILEC and CLEC proposals further illustrate this dilemma. The CLECs propose that benchmarks be strictly followed, causing instances where the ILECs must have 100 percent performance even though the parties have agreed that lower percentages allow a meaningful opportunity to compete. In compromise, the CLECs have offered small sample tables that allow performance thresholds to drop below the benchmark. However, these tables still may require overall performance levels to be well above the benchmark to avoid performance failure identification. For example, the CLEC-proposed small sample table for the 90 percent benchmark implies that the time criterion must be met

95.1 or 97.7 percent of the time for the ILEC to pass the measure– under conditions where the underlying process actually passes the benchmark.[2]

In contrast, the ILEC proposal for the 90 percent benchmark implies that the time criterion must be met 92.0 or 92.9 percent of the time for the ILEC to pass the measure[3] – again under conditions where the underlying process actually passes the benchmark. However both the ILEC and CLEC proposals' net result is to lower the effective benchmark performance level. Staff determined that the average effective level for the CLEC table was 89 percent and for the ILEC table was 83 percent.[4] For results with 100 or less orders, 13 and 27 percent of the results had effective benchmark percentages below 80 percent for the CLEC and ILEC small sample adjustment tables, respectively. The differences between CLEC and ILEC-proposed tables for the 95 and 99 percent benchmark parallel these differences.

There are a few aspects of small sample implementation that may allow maximizing the goals of both the ILECs and the CLECs, rather than trading the interests of one for the other. Staff examined four approaches:

---

[2] These percentages acknowledge random variation and assume a one-percent failure rate. The CLECs preferred using a derivation sample size of 20, which implies performance of 97.7 percent. However, the CLECs offered to compromise at a sample size of 100, which implies performance of 95.1 percent. Staff asked Pacific's Dr. Gleason for a copy of the MathCad© worksheet to calculate these implied performance levels. The worksheet is included as Attachment 1 to this appendix.

[3] This result also assumes only a one percent failure rate. Pacific prefers derivation sample sizes of 1000, which implies performance of 92.0 percent, but offers to use a derivation sample size of 400, which implies performance of 92.9 percent.

[4] These figures were calculated in several steps. First the effective performance level was calculated for each result in each table. For example, if for a sample of 4, one miss was allowed, the effective allowable performance level is 75 percent. Second, the percentage of sample sizes was calculated from January through May, 2000, data, for sample sizes of 100 or less (66% of total results). For example, samples with 2 orders accounted for 6.4 percent of these benchmark samples. Third, the effective percentage levels were weighted by these percentages. For example, since 7-order samples accounted for 3.2 percent of the results and 2-order samples accounted for 6.4 percent of the results, the 5-order samples' effective level was weighted twice the 7-order samples' level. And fourth, the overall average effective percentage was calculated from the weighted percentage. Using the same example, the 5-order samples' level essentially would be "counted" twice in determining the overall effective average compared to the 7-order samples' level. Example calculations were presented in Table 1.

(1) Application sample sizes currently are proposed to be the same for all benchmark percentage levels, and instead can be set to more closely fit the qualities of each different benchmark.
(2) Implied performance levels currently are proposed to require different degrees of improvement for different benchmarks; they can be made uniform by raising low improvement requirements and lowering high improvement requirements.
(3) Underlying process information: Aggregate CLEC results provide some indication of whether the underlying processes are passing or failing the benchmark, and thus can be used to guide a more targeted and appropriate application of the small sample tables.
(4) Small sample aggregation can alleviate the worst data "granularity" problems.

Application sample sizes

Applying small sample adjustment tables to the same sample size range for different benchmark percentage levels results in disparate treatment of the same problem. The "granularity" problem is of a different magnitude for different benchmark levels. For example, the sample size where a single delinquent order, or "miss," results in performance matching the nominal benchmark is very different for different benchmarks. For the 90 percent benchmark the sample size is 10. That is, one miss in ten equals 90 percent. In contrast, one miss in one hundred orders matches the 99 percent benchmark. Parallel to this phenomenon is the fact that the 99 percent benchmark experiences proportionately equal inaccuracies with much larger sample sizes than the 90 percent benchmark experiences. Or stated inversely, if applied to the same sample size range, the 99 percent benchmark experiences proportionately larger inaccuracies than the 90 percent benchmark. This results in a logical inconsistency where if the tolerances for one benchmark are optimized, then the other benchmarks are not optimal.

The solution described here selects reasonable tolerances and applies those tolerances consistently to all three benchmark percentage levels. Complete descriptions of the tolerances and methods to construct sample sizes is included as Attachment 2. Generally speaking, application sample sizes were selected that allowed discrepancies from the nominal benchmark percentage no greater than 10 percent of the allowable failure percentage.[5] This method resulted in sample sizes of 50, 100, and 500 for the 90, 95, and 99 percent benchmarks, respectively.[6]

Implied performance levels
Consistent implied performance criteria were used to select derivation sample sizes. First, the lower limit for each derivation sample size was determined by calculating the sample size that would allow no higher performance level than half the difference between the benchmark and 100 percent performance. For example, for the 90 percent benchmark, a derivation sample size of 125 implies a performance level of 95 percent, which is an increment of 5 percent or half the 10 percent allowed failures. Sample sizes were further adjusted by ensuring no effective adjusted percentage would be greater than the nominal benchmark percentage. This resulted in derivation sample sizes of 150, 300, and 1500 for the 90, 95, and 99 percent benchmarks, respectively. Attachment 2 lists the exact methods used.

Underlying process information
Application of small sample adjustment tables reduces what is analogous to a Type I error. That is, when an OSS process provides the service it should provide as defined by the benchmark, small sample adjustment tables reduce the likelihood of identifying spurious failures. On the other hand, adjustment tables increase the likelihood of what is analogous to a Type II error. That is, when an OSS process provides the service at levels lower than it should, small sample adjustment tables increase the likelihood that failures will not be detected.

Accuracy in benchmark decisions can be increased by applying small sample adjustment tables to situations where “Type I” errors are likely,

---

[5] For example, the 90 percent benchmark allows 10-percent failures. Ten percent of that failure allowance is one percent. Thus a small sample table is applied to all ranges of sample sizes where the average effective percentage is greater than 91 percent. Similarly, the 99 percent benchmark allows a one- percent failure. Ten percent of that failure allowance is one tenth of one percent. Thus a small sample table is applied to all ranges of sample sizes where the average effective percentage is greater than 99.1 percent. See Attachment 2 for the exact method used.

[6] Staff understands that pending Commission approval, only benchmark percentages of 90, 95, and 99 will be used in the final performance incentive plan. If it happens that other percentage levels are ultimately used, the method described in this appendix can be used to create any new sample size adjustment table.

and not to situations where "Type II" errors are likely. Analysis at the industry-wide aggregate performance level provides reasonably sufficient information regarding whether "Type I" or "Type II" errors are likely. If the aggregate performance level passes the benchmark, then CLEC-level performance that fails the benchmark has a greater likelihood of being a "Type I" error than if the aggregate performance failed the benchmark. And conversely, if the aggregate performance level fails the benchmark, then CLEC-level performance that passes the benchmark has a greater likelihood of being a Type II error than if the aggregate performance passed the benchmark. Therefore, a two-step small sample adjustment table application will maximize accuracy. First, since the large aggregate sample sizes provide a reasonable estimate of the process performance, they can be used to categorize performance results into "Type I error likely" and "Type II error likely" categories. For those samples where the aggregate performance passes ("Type I error likely"), small sample adjustment tables will be applied. For those samples where the aggregate performance fails ("Type II error likely"), small sample adjustment tables will not be applied – the nominal benchmark percentage will be the "pass / fail" criterion. Small sample adjustment tables will be used for the industry aggregate evaluation, however, in case any of these aggregate performance samples are small. The benefit of assessing and using the two categories to determine table application is best illustrated in the comparison of aggregate versus CLEC-specific sample sizes for the 90 and 99 percent benchmarks.[7] Table 2 shows the differences between aggregate and CLEC-specific result samples sizes. Aggregate sample sizes are typically large enough to assess whether "Type I" or "Type II" analogous error is likely for the much smaller CLEC-specific samples.

Table 2

Median sample sizes		
	Benchmark	
Sample type	90%	99%
CLEC-specific	36	5
Aggregate	9246	40725

[7] Sample sizes for the 95 percent benchmark are considerable larger. Only 14.7 percent of the CLEC-specific samples sizes are less than 100, and 14 percent of the aggregate sample sizes are less than approximately 9500.

Small sample aggregation

Table 1 illustrates that the granularity problem is most severe for the smallest samples. The most dramatic example is for the sample size of one. The only choices for an effective percentage level are zero and 100 percent. If the "ones" are aggregated, for example, into aggregates of five, then the choices are far better - 80 and 100 percent. Since aggregation of the smallest samples can alleviate the worst instances of the small sample problem, and since the equitable allocation of incentive payments still can be addressed in the incentive development phase of this proceeding, using the same aggregation rules as used for average-based parity measures can alleviate problems without unreasonably disadvantaging any party.

Methods

Based on the above principles, and using a MathCad© worksheet,[8] staff created small sample tables included here as Attachment 4. These tables are based on the following principles: (1) Application and derivation sample sizes should be set according to consistent relationships to the benchmark, (2) Implied performance should be no more than halfway between the benchmark and 100 percent performance.

Uniform criteria were applied to the three nominal benchmarks to establish the application sample size. These tables also provide a uniform limit of implied performance across the different benchmarks. They were constructed so that the implied performance would not exceed the midpoint between the benchmark nominal percentage and 100 performance. To accomplish this, the derivation sample sizes were 150, 300, and 1500 for the 90, 95, and 99 percent nominal benchmarks, respectively. With the midpoint expressed as 50 percent of the difference between the nominal percentage and 100 percent performance, the

---

[8] Staff requested Pacific Bell's consultant, Dr. Gleason, to provide staff with a copy of a program that would calculate the permitted misses for benchmarks in the form of a "small sample adjustment table." The Mathcad© program created by Dr. Gleason is included as Attachment 3. It is staff's understanding that AT&T's Dr. Mallows and Pacific's Dr. Gleason agreed on the methodology that Dr. Gleason subsequently forwarded to staff in the form of the MathCad© worksheets included here as Attachment 1 and 3. The worksheet in Attachment 1 calculates the implied performance level of different derivation sample sizes for different benchmark percentage levels. Staff used the worksheet to determine the sample sizes that would produce the desired performance level. Using the worksheet in Attachment 3, staff constructed the adjustment tables based on those sample sizes.

resultant implied performance difference was 44.2, 44.8, and 45.3 of the difference for the 90, 95, and 99 percent benchmarks, respectively.[9]

Summary

The small sample adjustment tables presented in Attachment 4 will be used in the following steps:

1. The number of performance "misses" for the CLEC industry-wide aggregate for each remedy plan benchmark submeasure will be compared to the number of permitted misses for all sample sizes covered by the related adjustment table. Industry aggregate performance will be identified as passing if the number of actual misses is less than or equal to the number of permitted misses, and identified as failing if otherwise.

2. For CLEC industry-wide aggregate sample sizes not covered by the related adjustment table, the actual performance percentage result will be compared to the benchmark nominal percentage value. Industry aggregate performance will be identified as passing if the actual performance percentage result is greater than or equal to the benchmark nominal percentage value, and identified as failing if otherwise.

3. For CLEC-specific analysis, results with sample sizes of four or less will be aggregated into a "small sample CLEC aggregate" for each submeasure. Each small sample CLEC aggregate performance result and all remaining non-aggregated CLEC performance results will be assessed.

4. For each submeasure where the CLEC industry-wide aggregate performance *fails* the benchmark, the actual performance percentage result for each small sample CLEC aggregate and each remaining non-aggregated CLEC result will be compared to the benchmark nominal

[9] In contrast, for example, the CLECs' percentage differences for the 3 benchmarks differed widely: 51.0, 63.7, and 85.1 for the 90, 95, and 99 percent benchmarks, respectively, using the derivation sample size of 100. Pacific's percentage differences for the 3 benchmarks also differed widely: 29.9, 40.3, and 67.9 for the 90, 95, and 99 percent benchmarks, respectively, using the derivation sample size of 400.

percentage value. Each individual or aggregate performance result will be identified as passing if the actual performance percentage result is greater than or equal to the benchmark nominal percentage value, and identified as failing if otherwise.

5. For sample sizes *covered* by the related adjustment table where the CLEC industry-wide aggregate performance *passes* the benchmark, the following shall apply for each submeasure. For each benchmark submeasure, the number of performance “misses” for each small sample CLEC aggregate and each remaining non-aggregated CLEC will be compared to the number of permitted misses. CLEC performance will be identified as passing if the number of actual misses is less than or equal to the number of permitted misses, and identified as failing if otherwise.

6. For sample sizes *not covered* by the related adjustment table where the CLEC industry-wide aggregate performance *passes* the benchmark, the following shall apply. The actual performance percentage result for each small sample CLEC aggregate and each remaining non-aggregated CLEC result will be compared to the benchmark nominal percentage value. Each individual or aggregate performance result will be identified as passing if the actual performance percentage result is greater than or equal to the benchmark nominal percentage value, and identified as failing if otherwise.

Mathcad worksheet: Small Sample Implied Performance Levels

Benchmarks for evaluation

$B := (.9 \quad .95 \quad .97 \quad .99)^T$

Reference sample sizes

$N := (150 \quad 300 \quad 500 \quad 1500)^T$

$i := 0 .. \text{length}(N) - 1 \qquad\qquad j := 0 .. \text{length}(B) - 1$

The following matrix gives the minimum number of "hits" consistent with the benchmark.

$$M_{i,j} := \text{ceil}\left(B_j \cdot N_i\right) \qquad M = \begin{bmatrix} 135 & 143 & 146 & 149 \\ 270 & 285 & 291 & 297 \\ 450 & 475 & 485 & 495 \\ 1.35 \cdot 10^3 & 1.425 \cdot 10^3 & 1.455 \cdot 10^3 & 1.485 \cdot 10^3 \end{bmatrix}$$

The following solve block calculates the performance level that meets the conditions that the failure rate on the benchmark given the sample size and the performance level should be 1%.

$p_j := \dfrac{B_j + 1}{2}$ Intial guesses for the solve block

Given

$\text{pbinom}(M - 1, N, p) = .01$

$f(N, M, p) := \text{Find}(p)$

$X_{i,j} := f\left(N_i, M_{i,j}, p_j\right)$

The matrix X has the performance levels by sample size and benchmark

$$X = \begin{bmatrix} 0.94416 & 0.98036 & 0.99139 & 0.99901 \\ 0.93371 & 0.97241 & 0.98612 & 0.99725 \\ 0.92721 & 0.96846 & 0.98352 & 0.99642 \\ 0.91666 & 0.96164 & 0.97877 & 0.99453 \end{bmatrix}$$

$Y := \text{augment}(N, X)$

$y := \text{augment}\left(0, B^T\right)$

$Z := \text{stack}(y, Y)$

The results are placed in a spreadsheet.

Z^T

Small sample adjustment table construction method.

To determine application sample sizes:

1. Calculate the net effective percentages for sample sizes without adjustments (i.e., use absolute application of benchmark percentage cutoffs).

2. Determine the table sub-ranges that are bounded by the different values where integer failures equal the benchmark. For example, for the 90-percent benchmark, the first integer failure results in a performance level equal to the benchmark at a sample size of 10. One (1) failure out of ten represents 90% performance. Thus the sub-ranges are 1-10, 21-30, 31-40, etc. For the 0.95 benchmark, the sub-range boundaries are 1-20, 21-40, 41-60, etc; and for the 0.99 benchmark the sub-range boundaries are 1-100, 101-200, 201-300, etc.

3. For each sub-range, exclude the sample size where the net effective percentage equals the benchmark and determine the average net effective percentage. For example, samples sizes of 10, 20, 30, etc., for the 90 percent benchmark are excluded. The application table will include all sub-ranges where the average effective percentage is greater than nominal benchmark value by 10 percent of the allowable “missed” percentage. For example, the 90-percent benchmark allows 10 percent “misses.” Ten percent of the allowable misses is one (1) percent. Therefore the table would include all sub-ranges where the average effective percentage is greater than 91 percent. The corresponding values for the 95 and 99-percent benchmarks are 95.5 and 99.1 percent, respectively. Following these criteria, adjustment tables will be applied to sample sizes of 50, 100, and 500 for the 90, 95, and 99 percent benchmarks, respectively.

Derivation sample sizes.

1. Smaller derivation sample sizes result in higher implied performance thresholds. Determine the lower bound for the derivation sample size by calculating adjustment tables that result in an implied performance value of no more than the midpoint of the interval between the nominal benchmark and perfect performance, 100 percent. For example, the midpoint between the 90-percent benchmark and 100 percent is 95 percent. Thus, the implied performance limits for adjustment tables are 95 percent for the 90 percent benchmark, 97.5 percent for the 95 percent benchmark, and 99.5 percent for the 99 percent benchmark.

2. Calculate the net effective benchmark percentages using adjustment tables to find the derivation sample size that is equal to or less than the limit. These values are 125, 246, and 1222 for the 90, 95, and 99 percent benchmarks, respectively.

3. If any net effective benchmark percentage is greater than the nominal benchmark percentage, then increase the derivation sample size until the net percentage is greater than the nominal percentage by no more than 10 percent of the allowable failure percentage. These individual sample size net effective percentage limits are 91, 95.5, and 99.1 percent for the 90, 95, and 99 percent nominal percentage benchmarks, respectively. Using these limits, the resulting derivation sample sizes are 150, 300, and 1500 for the 90, 95, and 99 percent nominal percentage benchmarks, respectively.

Mathcad worksheet: Creates small sample adjustment tables for different percentage benchmarks.

Set benchmarks for analysis.

$B := (.9 \quad .95 \quad .97 \quad .98 \quad .99 \quad .9925)^T$

Set reference sample size.

$N := 300$

Set probability of failing the benchmark at the reference sample size.

$P := .01$

Set probability of failing the benchmark with small samples.

$T1E := .1$

Set length of the small sample table.

$L := 100$

$J := \text{length}(B)$ $\qquad j := 0 .. J - 1$

The vector "b" gives the minimum number of successes permitted by the benchmark at the reference sample size.

$b_j := \text{ceil}\left(B_j \cdot N\right)$ $\qquad b^T = (270 \quad 285 \quad 291 \quad 294 \quad 297 \quad 298)$

The vector "p" gives initial guesses at the required performance levels

$$p_j := \frac{B_j + 1}{2}$$

The following function calculates performance levels that are consistent with the reference sample size N and probability P.

Given

$\text{pbinom}(b - 1, N, p) = P$

$f(b, p) := \mathrm{Find}(p)$

These are the required performance levels.

$F_j := f(b_j, p_j)$

$F^T = (0.9337126 \quad 0.9724129 \quad 0.9861193 \quad 0.9921846 \quad 0.9972458 \quad 0.9985427)$

Calculate the minimum number of misses for which the cumulative probability is less than T1E.

$$\mathrm{select}(n, P) := \left| \begin{array}{l} k \leftarrow 1 \\ \text{while} \quad \mathrm{pbinom}(n - k, n, P) \geq T1E \\ \quad k \leftarrow k + 1 \\ \text{return } k - 1 \end{array} \right.$$

$k := 0 .. L - 1$

$K_k := k + 1$

$Y_{k, j} := \mathrm{select}(K_k, F_j)$

$Z := \mathrm{augment}(K, Y)$

$\mathrm{head}_{j + 1} := B_j$

The following is the Small Sample Table.

$SST := \mathrm{stack}(\mathrm{head}^T, Z)$

Insert the Small Sample Table into an Excel spreadsheet

SST

Appendix K, Attachment 4

Benchmark Small Sample Adjustment Tables

90% Benchmark		95% Benchmark		99% Benchmark	
Sample size	Maximum permitted misses	Sample size	Maximum permitted misses	Sample size	Maximum permitted misses
1	0	1 to 3	0	1 to 19	0
2 to 9	1	4 to 19	1	20 to 97	1
10 to 20	2	20 to 40	2	98 to 202	2
21 to 31	3	41 to 63	3	203 to 319	3
32 to 44	4	64 to 88	4	320 to 445	4
45 to 50	5	89 to 100	5	446 to 500	5

Appendix L

List of Appearances

Respondents: Ed Kolto-Wininger and James B. Young, Attorneys at Law, for Pacific Bell; Marlin Ard and Elaine M. Duncan, Attorneys at Law, for Verizon California Inc.

Interested Parties: Evelyn C. Lee, Attorney at Law, for WorldCom, Inc.; Randolph Deutsch and Joseph Faber, Attorneys at Law, for AT&T Communications of California, Inc.; Richard L. Goldberg, Attorney at Law, for Sprint Communications Company LP; Theresa L. Cabral, Attorney at Law, for Mediaone Telecommunications of California and Karen Potkul, Attorney at Law, for XO, Inc. (formerly, Nextlink, Inc.)

Office of Ratepayer Advocates: Janice Grau, Attorney at Law.

(END OF APPENDIX L)